

JC20 Rec'd PCT/PTO SEP 28 2001

09/937779

TRANSMITTAL LETTER TO THE UNITED STATES

ATTORNEY'S DOCKET NUMBER 0093/000003

DESIGNATED/ELECTED OFFICE (DO/EO/US)  
CONCERNING A FILING UNDER 35 U.S.C. 371

U.S. APPLICATION NO. (If known, see 37 CFR 1.5)

INTERNATIONAL APPLICATION NO.  
PCT/EP 00/02701

INTERNATIONAL FILING DATE  
12 OCT 2000

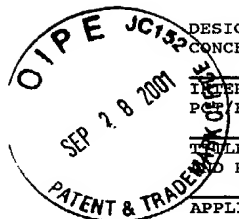
PRIORITY DATE CLAIMED  
March 28, 2000

TITLE OF INVENTION: A NEW CLASS OF ENZYMES IN THE BIOSYNTHETIC PATHWAY FOR THE PRODUCTION OF TRIACYLGLYCEROL AND RECOMBINANT DNA MOLECULES ENCODING THESE ENZYMES

APPLICANT(S) FOR DO/EO/US Anders DAHLQUIST; Ulf STAHL; Marit LENMAN; Antoni BANAS Hans RONNE; Sten STYMNE

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. /X/ This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.
  2. / / This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.
  3. /X/ This express request to begin national examination procedures (35 U.S.C.371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
  4. / / A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
  5. /X/ A copy of the International Application as filed (35 U.S.C. 371(c)(2)).
    - a./X/ is transmitted herewith (required only if not transmitted by the International Bureau).
    - b./ / has been transmitted by the International Bureau.
    - c./ / is not required, as the application was filed in the United States Receiving Office (RO/USO).
  6. /X/ A translation of the International Application into English (35 U.S.C. 371(c)(2)).
  7. / / Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3)).
    - a./ / are transmitted herewith (required only if not transmitted by the International Bureau).
    - b./ / have been transmitted by the International Bureau.
    - c./ / have not been made; however, the time limit for making such amendments has NOT expired.
    - d./ / have not been made and will not be made.
  8. / / A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
  9. /X/ An oath or declaration of the inventor(s) (35 U.S.C. 171(c)(4)).
  10. / / A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).
- Items 11. to 16. below concern other document(s) or information included:
- 11./X/ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
  12. / / An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
  - 13./X/ A FIRST preliminary amendment.  
/ / A SECOND or SUBSEQUENT preliminary amendment.
  14. / / A substitute specification.
  15. / / A change of power of attorney and/or address letter.
  - 16./X/ Other items or information.  
International Search Report  
International Preliminary Examination Report



JC05 Rec'd PCT/PTO 2 8 SEP 2001

09/937779

U.S. Appl. No. (If Known)	INTERNATIONAL APPLN. NO.	ATTORNEY'S DOCKET NO.
- - -	PCT/EP 00/02701	0093/000003

17. /X/ The following fees are submitted	<u>CALCULATIONS</u>	<u>PTO USE ONLY</u>
BASIC NATIONAL FEE (37 CFR 1.492(a)(1)-(5)):		
Search Report has been prepared by the		
EPO or JPO.....	\$860.00      860.00	
International preliminary examination fee paid to USPTO		
(37 CFR 1.482).....	\$690.00	
No international preliminary examination fee paid to		
USPTO (37 CFR 1.482) but international search fee paid		
to USPTO (37 CFR 1.445(a)(2)).....	\$710.00	
Neither international preliminary examination fee		
(37 CFR 1.482) nor international search fee		
(37 CFR 1.445(a)(2)) paid to USPTO .....	\$1000.00	
International preliminary examination fee paid to		
USPTO (37 CFR 1.482) and all claims satisfied pro		
-visions of PCT Article 33(2)-(4).....	\$100.00	
ENTER APPROPRIATE BASIC FEE AMOUNT =	\$      860.00	
Surcharge of \$130.00 for furnishing the oath or declaration		
later than      / / 20 / /30 months from the earliest		
claimed priority date (37 CFR 1.492(e)).		
<u>Claims</u>	<u>Number Filed</u>	<u>Number Extra</u>
Total Claims	22   -20	2
Indep. Claims	4   -3	1
Multiple dependent claim(s) (if applicable)		
		X\$18.      36.00
		X\$80.      80.00
		+270.
TOTAL OF ABOVE CALCULATION	=	116.00
Reduction of 1/4 for filing by small entity, if applicable.		
Verified Small Entity statement must also be filed		
(Note 37 CFR 1.9, 1.27, 1.28).		
SUBTOTAL	=	976.00
Processing fee of \$130. for furnishing the English		
translation later than / /20 / /30 months from the		
earliest claimed priority date (37 CFR 1.492(f)). +		
TOTAL NATIONAL FEE	=	976.00
Fee for recording the enclosed assignment (37 CFR 1.21(h)).		
The assignment must be accompanied by an appropriate cover		
sheet (37 CFR 3.28, 3.31) \$40.00 per property	=	
TOTAL FEES ENCLOSED	= \$	976.00
Amount to be		
refunded:	\$	
Charged	\$	

a./X/ A check in the amount of \$976.00 to cover the above fees is enclosed.

b./ / Please charge my Deposit Account No. \_\_\_\_\_ in the amount of \$ \_\_\_\_\_ to cover the above fees. A duplicate copy of this sheet is enclosed.

c./X/ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 11-0345. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

KEIL & WEINKAUF  
1101 Connecticut Ave., N.W.  
Washington, D. C. 20036

*Herbert B. Keil*  
\_\_\_\_\_  
Herbert B. Keil  
NAME  
18,967  
Registration No.

SIGNATURE

09/937779  
JC05 Rec'd PCT/PTO 2 8 SEP 2001

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of )  
DAHLQUIST et al. )  
PCT/EP 00/02701 )  
Intl. Filing Date: March 28, 2000 )  
US Serial No.: TO BE ASSIGNED )

BOX PCT

For: A NEW CLASS OF ENZYMES IN THE BIOSYNTHETIC PATHWAY FOR THE  
PRODUCTION OF TRIACYLGLYCEROL AND RECOMBINANT DNA MOLECULES  
ENCODING THESE ENZYMES

Honorable Commissioner of  
Patents and Trademarks  
Washington, D.C. 20231

PRELIMINARY AMENDMENT

Sir:

Prior to examination of the above-identified U.S. National  
Stage application, kindly amend the application as follows.

CLEAN VERSION OF ALL CLAIMS

Cancel claims 1-27, all the claims in this case, and substitute new claims 28-49 as follows:

28. An enzyme, designated as phospholipid:diacylglycerol acyltransferase (PDAT), catalyzing in an acyl-CoA-independent reaction the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol and comprising an amino acid sequence as set forth in SEQ ID NO:2 or a functional fragment, derivate, allele, homolog or isoenzyme thereof.

29. An enzyme according to claim 28, comprising an amino acid sequence encoded through a nucleotide sequence as set forth in SEQ ID NO:1 or a homologous nucleotide sequence which is at least about 40% identical to a nucleotide sequence of SEQ ID NO. 1.

30. An enzyme according to claim 28, comprising an amino acid sequence as set forth in SEQ ID NO. 16, 20 or 22, or a functional fragment, derivate, allele, homolog or isoenzyme thereof.

31. An enzyme according to claim 28, comprising an amino acid sequence selected from the group consisting of sequences as set forth in SEQ ID NO. 6, 8, 13, 14, 15, 17, 18, 25 and 27, a functional fragment, derivate, allele, homolog or isoenzyme thereof.

32. An enzyme according to claim 28, comprising an amino acid sequence encoded through a nucleotide sequence, a portion, derivate, allele or homolog thereof selected from the group consisting of sequences as set forth in SEQ ID NO. 1, 3, 4, 5, 7, 9, 10, 11, 12, 19, 21, 23, 24, 25, 26, 28, 29, 30 and 31, or a functional fragment, derivate, allele, homolog or isoenzyme of the enzyme encoding amino acid sequence.

33. A nucleotide sequence according to claim 32, selected



from the group consisting of sequences as set forth in SEQ ID NO. 1, 3, 4, 10, 11, 19, 21, 23, 24, 29 and 30, or a portion, derivate, allele or homolog thereof.

34. A partial nucleotide sequence corresponding to a full length nucleotide sequence according to claim 32, selected from the group consisting of sequences as set forth in SEQ ID NO. 5, 7, 9, 12, 25, 26, 28 and 31, or a portion, derivate, allele or homolog thereof.

35. A nucleotide sequence according to claim 32, comprising a nucleotide sequence which is at least 40% identical to a nucleotide sequence selected from the group consisting of those sequences set forth in SEQ ID NO. 1, 3, 4, 5, 7, 9, 10, 11, 12, 19, 21, 23, 24, 25, 26, 28, 29, 30 and 31.

36. A gene construct comprising a nucleotide sequence as set forth in SEQ ID No. 1, or a homologous nucleotide sequence which is at least about 40% identical to the nucleotide sequence of SEQ ID No. 1, which is operably linked to a heterologous nucleic.

37. A vector comprising a gene construct according to claim 36, or the nucleotide sequence as set forth in SEQ ID No. 1, or a homologous nucleotide sequence which is at least about 40% identical to the nucleotide sequence of SEQ ID No. 1.

38. A vector according to claim 37, which is an expression vector.

39. A vector according to claim 37, further comprising a selectable marker gene and/or nucleotide sequences for the replication in a host cell or the integration into the genome of the host cell.

40. A transgenic cell or organism comprising one or more of the following:

a) a nucleotide sequence  $a_1$ ) to  $a_4$ ),

b) a gene construct  $b_1$ ), and

c) a vector  $c_1$ ),

wherein

- $a_1$ ) is a nucleotide sequence as set forth in SEQ ID NO. 1 or a homologous nucleotide sequence which is at least about 40% identical to a nucleotide sequence of SEQ ID NO. 1,
- $a_2$ ) is a nucleotide sequence, a portion, derivate, allele or homolog thereof selected from the group consisting of sequences as set forth in SEQ ID NO. 1, 3, 4, 5, 7, 9, 10, 11, 12, 19, 21, 23, 24, 25, 26, 28, 29, 30 and 31, or a functional fragment, derivate, allele, homolog or isoenzyme of the enzyme encoding amino acid sequence,
- $a_3$ ) is a partial nucleotide sequence which corresponds to a full length nucleotide sequence selected from the group consisting of sequences as set forth in SEQ ID NO. 5, 7, 9, 12, 25, 26, 28 or 31, or a portion, derivate, allele or homolog thereof;
- $a_4$ ) is a nucleotide sequence which is at least 40% identical to a nucleotide sequence selected from the group consisting of those sequences set forth in SEQ ID NO. 1, 3, 4, 5, 7, 9, 10, 11, 12, 19, 21, 23, 24, 25, 26, 28, 29, 30 and 31,
- $b_1$ ) is a gene construct comprising a nucleotide sequence  $a$ ), operably linked to a heterologous nucleic acid, and
- $c_1$ ) is a vector comprising a gene construct  $b_1$ ), or a nucleotide sequence  $a_1$ ).

41. A transgenic cell or organism according to claim 40, which is an eucaryotic cell or organism.

42. A transgenic cell or organism according to claim 40, which is a yeast cell or a plant cell or a plant.

43. A transgenic cell or organism according to claim 40 having an altered biosynthetic pathway for the production of

triacylglycerol, characterized by the prevention of accumulation of undesirable fatty acids, which are harmful if present in high amounts in membrane lipids.

44. A transgenic cell or organism according to claim 40 having an altered, increased oil content.

45. A transgenic cell or organism according to claim 40, wherein the activity of PDAT is altered, characterized by an alteration in gene expression, catalytic activity and/or regulation of activity of the enzyme.

46. A process for the production of triacylglycerol, comprising growing a transgenic cell or organism according to claim 40 under conditions whereby the said nucleotide sequence

- a<sub>1</sub>) is a nucleotide sequence as set forth in SEQ ID NO. 1 or a homologous nucleotide sequence which is at least about 40% identical to a nucleotide sequence of SEQ ID NO. 1,
  - a<sub>2</sub>) is a nucleotide sequence, a portion, derivate, allele or homolog thereof selected from the group consisting of sequences as set forth in SEQ ID NO. 1, 3, 4, 5, 7, 9, 10, 11, 12, 19, 21, 23, 24, 25, 26, 28, 29, 30 and 31, or a functional fragment, derivate, allele, homolog or isoenzyme of the enzyme encoding amino acid sequence,
  - a<sub>3</sub>) is a partial nucleotide sequence which corresponds to a full length nucleotide sequence selected from the group consisting of sequences as set forth in SEQ ID NO. 5, 7, 9, 12, 25, 26, 28 or 31, or a portion, derivate, allele or homolog thereof;
  - a<sub>4</sub>) is a nucleotide sequence which is at least 40% identical to a nucleotide sequence selected from the group consisting of those sequences set forth in SEQ ID NO. 1, 3, 4, 5, 7, 9, 10, 11, 12, 19, 21, 23, 24, 25, 26, 28, 29, 30 and 31,
- is expressed.

47. Triacylglycerols produced by a process according to claim 46.

48. Use of a nucleotide sequence selected from the group of  
a<sub>1</sub>) is a nucleotide sequence as set forth in SEQ ID NO. 1 or a homologous nucleotide sequence which is at least about 40% identical to a nucleotide sequence of SEQ ID NO. 1,

a<sub>2</sub>) is a nucleotide sequence, a portion, derivate, allele or homolog thereof selected from the group consisting of sequences as set forth in SEQ ID NO. 1, 3, 4, 5, 7, 9, 10, 11, 12, 19, 21, 23, 24, 25, 26, 28, 29, 30 and 31, or a functional fragment, derivate, allele, homolog or isoenzyme of the enzyme encoding amino acid sequence,

a<sub>3</sub>) is a partial nucleotide sequence which corresponds to a full length nucleotide sequence selected from the group consisting of sequences as set forth in SEQ ID NO. 5, 7, 9, 12, 25, 26, 28 or 31, or a portion, derivate, allele or homolog thereof;

a<sub>4</sub>) is a nucleotide sequence which is at least 40% identical to a nucleotide sequence selected from the group consisting of those sequences set forth in SEQ ID NO. 1, 3, 4, 5, 7, 9, 10, 11, 12, 19, 21, 23, 24, 25, 26, 28, 29, 30 and 31,

or an enzyme selected from the group of

d<sub>1</sub>) is an enzyme, designated as phospholipid:diacylglycerol acyltransferase (PDAT), catalyzing in an acyl-CoA-independent reaction the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol and comprising an amino acid sequence as set forth in SEQ ID NO. 2 or a functional fragment, derivate, allele, homolog or isoenzyme thereof,

d<sub>2</sub>) is an enzyme d), comprising an amino acid sequence as set forth in SEQ ID NO. 16, 20 or 22, or a functional fragment,

derivate, allele, homolog or isoenzyme thereof,  
d<sub>3</sub>) is an enzyme d<sub>1</sub>), comprising an amino acid sequence selected from the group consisting of sequences as set forth in SEQ ID NO. 6, 8, 13, 14, 15, 17, 18, 25 and 27, or a functional fragment, derivate, allele, homolog or isoenzyme thereof, for the production of triacylglycerol and/or triacylglycerols with uncommon fatty acids, comprising medium chain fatty acids, hydroxylated fatty acids, epoxygenated fatty acids and acetylenic fatty acids.

49. Use of a nucleotide sequence selected from the group of
- a<sub>1</sub>) is a nucleotide sequence as set forth in SEQ ID NO. 1 or a homologous nucleotide sequence which is at least about 40% identical to a nucleotide sequence of SEQ ID NO. 1,
  - a<sub>2</sub>) is a nucleotide sequence, a portion, derivate, allele or homolog thereof selected from the group consisting of sequences as set forth in SEQ ID NO. 1, 3, 4, 5, 7, 9, 10, 11, 12, 19, 21, 23, 24, 25, 26, 28, 29, 30 and 31, or a functional fragment, derivate, allele, homolog or isoenzyme of the enzyme encoding amino acid sequence,
  - a<sub>3</sub>) is a partial nucleotide sequence which corresponds to a full length nucleotide sequence selected from the group consisting of sequences as set forth in SEQ ID NO. 5, 7, 9, 12, 25, 26, 28 or 31, or a portion, derivate, allele or homolog thereof;
  - a<sub>4</sub>) is a nucleotide sequence which is at least 40% identical to a nucleotide sequence selected from the group consisting of those sequences set forth in SEQ ID NO. 1, 3, 4, 5, 7, 9, 10, 11, 12, 19, 21, 23, 24, 25, 26, 28, 29, 30 and 31,
- or an enzyme selected from the group of
- d<sub>1</sub>) is an enzyme, designated as phospholipid:diacylglycerol acyltransferase (PDAT), catalyzing in an acyl-CoA-independent

reaction the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol and comprising an amino acid sequence as set forth in SEQ ID NO. 2 or a functional fragment, derivate, allele, homolog or isoenzyme thereof,

d<sub>2</sub>) is an enzyme d<sub>1</sub>), comprising an amino acid sequence as set forth in SEQ ID NO. 16, 20 or 22, or a functional fragment, derivate, allele, homolog or isoenzyme thereof,

d<sub>3</sub>) is an enzyme d<sub>1</sub>), comprising an amino acid sequence selected from the group consisting of sequences as set forth in SEQ ID NO. 6, 8, 13, 14, 15, 17, 18, 25 and 27, or a functional fragment, derivate, allele, homolog or isoenzyme thereof,

for the transformation of any cell or organism in order to be expressed in this cell or organism and result in an altered, preferably increased oil content of this cell or organism.

REMARKS

The claims have been amended to eliminate multiple dependency and to place them in better form for U.S. practice. Further, amendments made in the international stage, albeit not under Article 19, are also incorporated. The changes made in the claims were as follows:

- i. the subject-matter of claim 2 and 3 was included in claim 1 (now claim 28);
- ii. claim 29 was introduced according to the disclosed homologous nucleotide sequence on page 6 of the specification;
- iii. claims 7, 8 and 25 were deleted;
- iv. claims 30-42 essentially correspond to claims 4-6 and 9-18 of the international application;
- v. claim 23 was incorporated into claim 19, which is now claim 43;
- vi. claim 20, now 44, was amended by defining the altered oil content according to page 9, line 9;
- vii. claim 22 was incorporated into claim 21, which is now claim 45;
- viii. claim 46 essentially corresponds to claim 24 of the international application.
- ix. in claim 26, now 47, triacylglycerols with uncommon fatty acids were defined according to page 10, line 29;

DAHLQUIST et al. et al., oz 0093/000003

x. claim 48 essentially corresponds to claim 27 of the international application;

xi. in the other claims, editorial amendments were made.

Favorable action on the application is solicited.

Respectfully submitted,

KEIL & WEINKAUF



Herbert B. Keil  
Reg. No. 18,967

1101 Connecticut Avenue, N.W.  
Washington, D.C. 20036  
(202) 659-0100  
HBK/kas



JC09 Rec'd PCT/PTO 07 JUN 2002



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Re Application of

DAHLQUIST et al.

Serial No. 09/937,779

Filed: September 28, 2001

For: NEW CLASS OF ENZYMES IN THE BIOSYNTHETIC PATHWAY FOR THE PRODUCTION OF TRIACYCLOGLYCEROL AND RECOMBINANT DNA MOLECULES ENCODING THESE ENZYMES

BOX PCT

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to Commissioner of Patents and Trademarks, Washington, D.C. 20231, on:

June 5, 2002

Date of Deposit Karen Stamper

Person Making Deposit

Signature

June 5, 2002

Date of Signature

Honorable Commissioner of  
Patents and Trademarks  
Washington, D.C. 20231

PRELIMINARY AMENDMENT  
AND  
RESPONSE TO NOTIFICATION OF MISSING REQUIREMENTS UNDER 35 USC 371

Sir:

In response to the Notification of Missing Requirements under 35 USC 371, a copy of the Sequence Listing in computer readable form is attached hereto. The content of the paper copy of the Sequence Listing and the copy of the Sequence Listing in computer readable form is the same, and includes no new matter.

IN THE SPECIFICATION

Delete the sequence listing in the specification on pages 1/58 to 58/58 and substitute with the attached replacement sequence listing on separate pages 1-52.

**REMARKS**

It is believed that by submitting the present amendment and sequence listing diskette, the application now fully complies with the requirements of 37 CFR 1.821-1.825. Favorable action by the examiner is solicited.

Please charge any shortage in fees due in connection with the filing of this paper, including Extension of Time fees to Deposit Account No. 11-0345. Please credit any excess fees to such account.

Respectfully submitted,

KEIL & WEINKAUF

06/13/2002 HKAYPAGH 00000090 110345 09937779  
01 FC:154 130.00 CH



Herbert B. Keil  
Reg. No. 18,967

1350 Connecticut Ave., N.W.  
Washington, D.C. 20036  
(202)659-0100

HBK/DSK/kas

06/13/2002 HKAYPAGH 00000090 110345 09937779  
02 FC:965 84.00 CH

A NEW CLASS OF ENZYMES IN THE BIOSYNTHETIC PATHWAY FOR THE  
PRODUCTION OF TRIACYLGLYCEROL AND RECOMBINANT DNA  
MOLECULES ENCODING THESE ENZYMES

- 5 The present invention relates to the isolation, identification and characterization of recombinant DNA molecules encoding enzymes catalysing the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol.
- 10 Triacylglycerol (TAG) is the most common lipid-based energy reserve in nature. The main pathway for synthesis of TAG is believed to involve three sequential acyl-transfers from acyl-CoA to a glycerol backbone (1, 2). For many years, acyl-CoA : diacylglycerol acyltransferase (DAGAT), which catalyzes the third acyl transfer reaction, was thought to be the only unique enzyme involved in
- 15 TAG synthesis. It acts by diverting diacylglycerol (DAG) from membrane lipid synthesis into TAG (2). Genes encoding this enzyme were recently identified both in the mouse (3) and in plants (4, 5), and the encoded proteins were shown to be homologous to acyl-CoA : cholesterol acyltransferase (ACAT). It was also recently reported that another DAGAT exists in the oleaginous fungus
- 20 *Mortierella ramanniana*, which is unrelated to the mouse DAGAT, the ACAT gene family or to any other known gene (6).

The instant invention relates to novel type of enzymes and their encoding genes for transformation. More specifically, the invention relates to use of a

25 type of genes encoding a not previously described type of enzymes hereinafter designated phospholipid:diacylglycerol acyltransferases (PDAT), whereby this enzyme catalyses an acyl-CoA-independent reaction. The said type of genes expressed alone in transgenic organisms will enhance the total amount of oil (triacylglycerols) produced in the cells. The PDAT genes, in combination with a

30 gene for the synthesis of an uncommon fatty acid will, when expressed in transgenic organisms, enhance the levels of the uncommon fatty acids in the triacylglycerols.

There is considerable interest world-wide in producing chemical feedstock, such as fatty acids, for industrial use from renewable plant resources rather than non-renewable petrochemicals. This concept has broad appeal to manufacturers and consumers on the basis of resource conservation and provides significant opportunity to develop new industrial crops for agriculture.


There is a diverse array of unusual fatty acids in oils from wild plant species and these have been well characterised. Many of these acids have industrial potential and this has led to interest in domesticating relevant plant species to enable agricultural production of particular fatty acids.

Development in genetic engineering technologies combined with greater understanding of the biosynthesis of unusual fatty acids now makes it possible to transfer genes coding for key enzymes involved in the synthesis of a particular fatty acid from a wild species into domesticated oilseed crops. In this way individual fatty acids can be produced in high purity and quantities at moderate costs.

In all crops like rape, sunflower, oilpalm etc., the oil (i.e. triacylglycerols) is the most valuable product of the seeds or fruits and other compounds like starch, protein, and fibre is regarded as by-products with less value. Enhancing the quantity of oil per weight basis at the expense of other compounds in oil crops would therefore increase the value of crop. If genes, regulating the allocation of reduced carbon into the production of oil can be up-regulated, the cells will accumulate more oil on the expense of other products. Such genes might not only be used in already high oil producing cells, such as oil crops, but could also induce significant oil production in moderate or low oil containing crops such as e.g. soy, oat, maize, potato, sugarbeets, and turnips as well as in micro-organisms.

30

Summary of the invention



Many of the unusual fatty acids of interest, e.g. medium chain fatty acids, hydroxy fatty acids, epoxy fatty acids and acetylenic fatty acids, have physical properties that are distinctly different from the common plant fatty acids. The present inventors have found that, in plant species naturally accumulating these uncommon fatty acids in their seed oil (i.e. triacylglycerol), these acids are absent, or present in very low amounts in the membrane (phospho)lipids of the seed. The low concentration of these acids in the membrane lipids is most likely a prerequisite for proper membrane function and thereby for proper cell functions. One aspect of the invention is that seeds of transgenic crops can be made to accumulate high amounts of uncommon fatty acids if these fatty acids are efficiently removed from the membrane lipids and channelled into seed triacylglycerols.

The inventors have identified a novel class of enzymes in plants catalysing the transfer of fatty acids from phospholipids to diacylglycerol in the production of triacylglycerol through an acyl-CoA-independent reaction and that these enzymes (phospholipid:diacylglycerol acyltransferases, abbreviated as PDAT) are involved in the removal of hydroxylated, epoxygenated fatty acids, and probably also other uncommon fatty acids such as medium chain fatty acids, from phospholipids in plants.

This enzyme reaction was shown to be present in microsomal preparations from baker's yeast (*Saccharomyces cerevisiae*). The instant invention further pertains to an enzyme comprising an amino acid sequence as set forth in SEQ ID No. 2 or a functional fragment, derivate, allele, homolog or isoenzyme thereof. A so called 'knock out' yeast mutant, disrupted in the respective gene was obtained and microsomal membranes from the mutant was shown to totally lack PDAT activity. Thus, it was proved that the disrupted gene encodes a PDAT enzyme (SEQ ID NO. 1 and 2). Furtherm, this PDAT enzyme is characterized through the amino acid sequence as set forth in SEQ ID NO 2 containing a lipase motif of the conserved sequence string FXKWVEA.

The instant invention pertains further to an enzyme comprising an amino acid sequence as set forth in SEQ ID NO. 16, 20 or 22 or a functional fragment, derivate, allele, homolog or isoenzyme thereof.

Further genes and/or proteins of so far unknown function were identified and are contemplated within the scope of the instant invention. A gene from *Schizosaccharomyces pombe*, SPBC776.14 (SEQ ID NO. 3), a putative open reading frame CAA22887 of the SPBC776.14 (SEQ ID NO. 13) were identified. Further *Arabidopsis thaliana* genomic sequences (SEQ ID NO. 4, 10 and 11) coding for putative proteins were identified, as well as a putative open reading frame AAC80628 from the *A. thaliana* locus AC 004557 (SEQ ID NO. 14) and a putative open reading frame AAD10668 from the *A. thaliana* locus AC 003027 (SEQ ID NO. 15) were identified.

Also, a partially sequenced cDNA clone from *Neurospora crassa* (SEQ ID NO. 9) and a *Zea mays* EST (Extended Sequence Tac) clone (SEQ ID NO. 7) and corresponding putative amino acid sequence (SEQ ID NO. 8) were identified. Finally, two cDNA clones were identified, one *Arabidopsis thaliana* EST (SEQ ID NO. 5 and corresponding predicted amino acid sequence SEQ ID NO. 6) and a *Lycopersicon esculentum* EST clone (SEQ ID NO. 12) were identified. Further, enzymes designated as PDAT comprising an amino acid sequence selected from the group consisting of sequences as set forth in SEQ ID NO 6, 17, 18, 25 or 27 containing a lipase motif FXKWVEA are contemplated within the scope of the invention. Moreover, an enzyme comprising an amino acid sequence encoded through a nucleotide sequence, a portion, derivate, allele or homolog thereof selected from the group consisting of sequences as set forth in SEQ ID No. 1, 3, 4, 5, 7, 9, 10, 11, 12, 19, 21, 23, 24, 25, 26, 28, 29, 30 or 31 or a functional fragment, derivate, allele, homolog or isoenzyme of the enzyme encoding amino acid sequence are included within the scope of the invention.

A functional fragment of the instant enzyme is understood to be any polypeptide sequence which shows specific enzyme activity of a phospholipid:diacylglycerol acyltransferase (PDAT). The length of the functional fragment can for example vary in a range from about  $660 \pm 10$  amino acids to

660  $\pm$  250 amino acids, preferably from about 660  $\pm$  50 to 660  $\pm$ 100 amino acids, whereby the „basic number“ of 660 amino acids corresponds in this case to the polypeptide chain of the PDAT enzyme of SEQ ID NO. 2 encoded by a nucleotide sequence according to SEQ ID NO. 1. Consequently, the „basic number“ of functional fulllength enzyme can vary in correspondance to the encoding nucleotide sequence.

A portion of the instant nucleotide sequence is meant to be any nucleotide sequence encoding a polypeptid which shows specific activity of a phospholipid:diacylglycerol acyltransferase (PDAT). The length of the nucleotide portion can vary in a wide range of about several hundreds of nucleotides based upon the coding region of the gene or a highly conserved sequence. For example the length varies in a range form about 1900  $\pm$  10 to 1900  $\pm$  1000 nucleotides, preferably form about 1900  $\pm$  50 to 1900  $\pm$ 700 and more preferably form about 1900  $\pm$  100 to 1900  $\pm$  500 nucleotides. whereby the „basic number“ of 1900 nucleotides corresponds in this case to the encoding nucleotide sequence of the PDAT enzyme of SEQ ID NO. 1. Consequently, the „basic number“ of functional fulllength gene can vary.

An allelic variant of the instant nucleotide sequence is understood to be any different nucleotide sequence which encodes a polypeptide with a functionally equivalent function. The alleles pertain naturally occurring variants of the instant nucleotide sequences as well as synthetic nucleotide sequences produced by methods known in the art. Contemplated are even altered nucleotide sequences which result in an enzyme with altered activity and/or regulation or which is resistant against specific inhibitors. The instant invention further includes natural or synthetic mutations of the originally isolated nucleotide sequences. These mutations can be substitution, addition, deletion, inversion or insertion of one or more nucleotides.

A homologous nucleotide sequence is understood to be a complementary sequence and/or a sequence which specifically hybridizes with the instant nucleotide sequence. Hybridizing sequences include similar sequences selected from the group of DNA or RNA which specifically interact to the instant

nucleotide sequences under at least moderate stringency conditions which are known in the art. A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-  
5 65°C. This further includes short nucleotide sequences of e.g. 10 to 30 nucleotides, preferably 12 to 15 nucleotides. Included are also primer or hybridization probes.

A homologous nucleotide sequence included within the scope of the instant  
10 invention is a sequence which is at least about 40%, preferably at least about 50 % or 60%, and more preferably at least about 70%, 80% or 90% and most preferably at least about 95%, 96%, 97%, 98% or 99% or more homologous to a nucleotide sequence of SEQ ID NO. 1.

All of the aforementioned definitions are true for amino acid sequences and  
15 functional enzymes and can easily transferred by a person skilled in the art.

Isoenzymes are understood to be enzymes which have the same or a similar substrate specificity and/or catalytic activity but a different primary structure.

20 In a first embodiment, this invention is directed to nucleic acid sequences that encode a PDAT. This includes sequences that encode biologically active PDATs as well as sequences that are to be used as probes, vectors for transformation or cloning intermediates. The PDAT encoding sequence may encode a complete or partial sequence depending upon the intended use. All or  
25 a portion of the genomic sequence, cDNA sequence, precursor PDAT or mature PDAT is intended.

Further included is a nucleotide sequence selected from the group consisting of sequences set forth in SEQ ID No. 1, 3, 4, 10, 11, 19, 21, 23, 24, 29 or 30 or a  
30 portion, derivate, allele or homolog thereof. The invention pertains a partial nucleotide sequence corresponding to a fulllength nucleotide sequence selected from the group consisting of sequences set forth in SEQ ID No. 5, 7, 9, 12, 25, 26, 28 or 31 or a portion, derivate, allele or homolog thereof. Moreover, a



nucleotide sequence comprising a nucleotide sequence which is at least 40% homologous to a nucleotide sequence selected from the group consisting of those sequences set forth in SEQ ID No. 1, 3, 4, 5, 7, 9, 10, 11, 12, 19, 21, 23, 24, 25, 26, 28, 29, 30 or 31 is contemplated within the scope of the invention.

5

The instant invention pertains to a gene construct comprising a said nucleotide sequences of the instant invention which is operably linked to a heterologous nucleic acid.

The term operably linked means a serial organisation e.g. of a promotor, coding  
10 sequence, terminator and/or further regulatory elements whereby each element can fulfill its original function during expression of the nucleotide sequence.

Further, a vector comprising of a said nucleotide sequence of the instant invention is contemplated in the instant invention. This includes also an  
15 expression vector as well as a vector further comprising a selectable marker gene and/or nucleotide sequences for the replication in a host cell and/or the integration into the genome of the host cell.

In a different aspect, this invention relates to a method for producing a PDAT in  
20 a host cell or progeny thereof, including genetically engineered oil seeds, yeast and moulds or any other oil accumulating organism, via the expression of a construct in the cell. Cells containing a PDAT as a result of the production of the PDAT encoding sequence are also contemplated within the scope of the invention.

25

Further, the invention pertains a transgenic cell or organism containing a said nucleotide sequence and/or a said gene construct and/or a said vector. The object of the instant invention is further a transgenic cell or organism which is an eucaryotic cell or organism. Preferably, the transgenic cell or organism is a  
30 yeast cell or a plant cell or a plant. The instant invention further pertains said transgenic cell or organism having an altered biosynthetic pathway for the production of triacylglycerol. A transgenic cell or organism having an altered oil content is also contemplated within the scope of this invention.

Further, the invention pertains a transgenic cell or organism wherein the activity of PDAT is altered in said cell or organism. This altered activity of PDAT is characterized by an alteration in gene expression, catalytic activity and/or regulation of activity of the enzyme. Moreover, a transgenic cell or organism is included in the instant invention, wherein the altered biosynthetic pathway for the production of triacylglycerol is characterized by the prevention of accumulation of undesirable fatty acids in the membrane lipids.

10 In a different embodiment, this invention also relates to methods of using a DNA sequence encoding a PDAT for increasing the oil-content within a cell.

Another aspect of the invention relates to the accommodation of high amounts of uncommon fatty acids in the triacylglycerol produced within a cell, by introducing a DNA sequence producing a PDAT that specifically removes these fatty acids from the membrane lipids of the cell and channel them into triacylglycerol. Plant cells having such a modification are also contemplated herein.

20 Further, the invention pertains a process for the production of triacylglycerol, comprising growing a said transgenic cell or organism under conditions whereby the said nucleotide sequence is expressed and whereby the said transgenic cells comprising a said enzyme catalysing the transfer of fatty acids from phospholipids to diacylglycerol forming triacylglycerol.

25 Moreover, triacylglycerols produced by the aforementioned process are included in scope of the instant invention.

30 Object of the instant invention is further the use of an instant nucleotide sequence and/or a said enzyme for the production of triacylglycerol and/or triacylglycerols with uncommon fatty acids. The use of a said instant nucleotide sequence and/or a said enzyme of the instant invention for the transformation of any cell or organism in order to be expressed in this cell or organism and

result in an altered, preferably increased oil content of this cell or organism is also contemplated within the scope of the instant invention.

A PDAT of this invention includes any sequence of amino acids, such as a  
5 protein, polypeptide or peptide fragment obtainable from a microorganism,  
animal or plant source that demonstrates the ability to catalyse the production  
of triacylglycerol from a phospholipid and diacylglycerol under enzyme reactive  
conditions. By „enzyme reactive conditions“ is meant that any necessary  
conditions are available in an environment (e.g., such factors as temperature,  
10 pH, lack of inhibiting substances) which will permit the enzyme to function.

Other PDATs are obtainable from the specific sequences provided herein.  
Furthermore, it will be apparent that one can obtain natural and synthetic  
PDATs, including modified amino acid sequences and starting materials for  
15 synthetic-protein modelling from the exemplified PDATs and from PDATs which  
are obtained through the use of such exemplified sequences. Modified amino  
acid sequences include sequences that have been mutated, truncated,  
increased and the like, whether such sequences were partially or wholly  
synthesised. Sequences that are actually purified from plant preparations or  
20 are identical or encode identical proteins thereto, regardless of the method  
used to obtain the protein or sequence, are equally considered naturally  
derived.

Further, the nucleic acid probes (DNA and RNA) of the present invention can  
25 be used to screen and recover „homologous“ or „related“ PDATs from a variety  
of plant and microbial sources.

Further, it is also apparent that a person skilled in the art can, with the  
information provided in this application, in any organism identify a PDAT  
30 activity, purify an enzyme with this activity and thereby identify a „non-  
homologous“ nucleic acid sequence encoding such an enzyme.

The present invention can be essentially characterized by the following aspects:

1. Use of a PDAT gene (genomic clone or cDNA) for transformation.
- 5 2. Use of a DNA molecule according to item 1 wherein said DNA is used for transformation of any organism in order to be expressed in this organism and result in an active recombinant PDAT enzyme in order to increase oil content of the organism.
3. Use of a DNA molecule of item 1 wherein said DNA is used for  
10 transformation of any organism in order to prevent the accumulation of undesirable fatty acids in the membrane lipids.
4. Use according to item 1, wherein said PDAT gene is used for transforming transgenic oil accumulating organisms engineered to produce any  
15 uncommon fatty acid which is harmful if present in high amounts in membrane lipids, such as medium chain fatty acids, hydroxylated fatty acids, epoxygenated fatty acids and acetylenic fatty acids.
5. Use according to item 1, wherein said PDAT gene is used for transforming organisms, and wherein said organisms are crossed with other oil  
20 accumulating organisms engineered to produce any uncommon fatty acid which is harmful if present in high amounts in membrane lipids, comprising medium chain fatty acids, hydroxylated fatty acids, epoxygenated fatty acids and acetylenic fatty acids.
6. Use according to item 1, wherein the enzyme encoded by said PDAT gene or cDNA is coding for a PDAT with distinct acyl specificity.
- 25 7. Use according to item 1 wherein said PDAT encoding gene or cDNA, is derived from *Saccharomyces cerevisiae*, or contain nucleotide sequences coding for an amino acid sequence 30% or more identical to the amino acid sequence of PDAT as presented in SEQ. ID. NO. 2.
8. Use according to item 1 wherein said PDAT encoding gene or cDNA is  
30 derived from *Saccharomyces cerevisiae*, or contain nucleotide sequences coding for an amino acid sequence 40% or more identical to the amino acid sequence of PDAT as presented in SEQ. ID. NO. 2.

9. Use according to item 1 wherein said PDAT encoding gene or cDNA is derived from *Saccharomyces cerevisiae*, or contain nucleotide sequences coding for an amino acid sequence 60% or more *identical* to the amino acid sequence of PDAT as presented in SEQ. ID. NO. 2.
- 5 10. Use according to item 1 wherein said PDAT encoding gene or cDNA is derived from *Saccharomyces cerevisiae*, or contain nucleotide sequences coding for an amino acid sequence 80% or more identical to the amino acid sequence of PDAT as presented in SEQ. ID. NO. 2.
- 10 11. Use according to item 1 wherein said PDAT encoding gene or cDNA is derived from plants or contain nucleotide sequences coding for an amino acid sequence 40% or more identical to the amino acid sequence of PDAT from *Arabidopsis thaliana* or to the protein encoded by the fulllength counterpart of the partial Zea mays, Lycopersicon esculentum, or Neurospora crassa cDNA clones.
- 15 12. Transgenic oil accumulating organisms comprising, in their genome, a PDAT gene transferred by recombinant DNA technology or somatic hybridization.
13. Transgenic oil accumulating organisms according to item 12 comprising, in their genome, a PDAT gene having specificity for substrates with a particular uncommon fatty acid and the gene for said uncommon fatty acid.
- 20 14. Transgenic organisms according to item 12 or 13 which are selected from the group consisting of fungi, plants and animals.
15. Transgenic organisms according to item 12 or 13 which are selected from the group of agricultural plants.
- 25 16. Transgenic organisms according to item 12 or 13 which are selected from the group of agricultural plants and where said PDAT gene is expressed under the control of a storage organ specific promotor.
17. Transgenic organisms according to item 12 or 13 which are selected from the group of agricultural plants and where said PDAT gene is expressed under the control of a seed promotor.
- 30 18. Oils from organisms according to item 12 – 17.
19. A method for altering acyl specificity of a PDAT by alteration of the nucleotide sequence of a naturally occurring encoding gene and as a

consequence of this alternation creating a gene encoding for an enzyme with novel acyl specificity.

20. A protein encoded by a DNA molecule according to item 1 or a functional fragment thereof.

5 21. A protein of item 20 designated phospholipid:diacylglycerol acyltransferase.

22. A protein of item 21 which has a distinct acyl specificity.

23. A protein of item 13 having the amino acid sequence as set forth in SEQ, ID NO. 2, 13, 14 or 15 (and the proteins encoded by the fulllength or partial genes set forth in SEQ. ID. NO. 1, 3, 4, 5, 7, 9, 10, 11 or 12) or an amino

10 acid sequence with at least 30 % homology to said amino acid sequence.

24. A protein of item 23 isolated from *Saccharomyces cerevisiae*.

15

20 General methods:

Yeast strains and plasmids. The wild type yeast strains used were either FY1679 (*MAT $\alpha$  his3- $\Delta$ 200 leu2- $\Delta$ 1 trp1- $\Delta$ 6 ura3-52*) or W303-1A (*MAT $\alpha$  ADE2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1*) (7). The YNR008w::KanMX2 disruption strain FVKT004-04C(AL), which is congenic to FY1679, was

25 obtained from the Euroscarf collection (8). A 2751 bp fragment containing the YNR008w gene with 583 bp of 5' and 183 bp of 3' flanking DNA was amplified from W303-1A genomic DNA using *Taq* polymerase with 5'-TCTCCATCTTCTGCAAAACCT-3' and 5'-CCTGTCAAAAACCTTCTCCTC-3' as primers. The resulting PCR product was purified by agarose gel electrophoresis

30 and cloned into the *EcoRV* site of pBluescript (pbluescript-pdat). For complementation experiments, the cloned fragment was released from pBluescript by *HindIII*-*SacI* digestion and then cloned between the *HindIII* and *SacI* sites of pFL39 (9), thus generating pUS1. For overexpression of the PDAT

gene, a 2202 bp *EcoRI* fragment from the pBluscript plasmid which contains only 24 bp of 5' flanking DNA was cloned into the *Bam*HI site of the *GAL1-TPK2* expression vector pJN92 (12), thus generating pUS4.

5        Microsomal preparations. Microsomes from developing seeds of sunflower (*Helianthus annuus*), *Ricinus communis* and *Crepis palaestina* were prepared using the procedure of Stobart and Stymne (11). To obtain yeast microsomes, 1g of yeast cells (fresh weight) was re-suspended in 8 ml of ice-cold buffer (20 mM Tris-Cl, pH 7.9, 10 mM MgCl<sub>2</sub>, 1 mM EDTA, 5 % (v/v) glycerol, 1 mM DTT,  
10    0.3 M ammonium sulfate) in a 12 ml glass tube. To this tube, 4 ml of glass beads (diameter 0.45-0.5 mm) were added, and the tube was then heavily shaken (3 x 60 s) in an MSK cell homogenizer (B. Braun Melsungen AG, Germany). The homogenized suspension was centrifuged at 20,000 x g for 15 min at 6°C and the resulting supernatant was again centrifuged at 100,000 x g  
15    for 2 h at 6°C. The 100,000 x g pellet was resuspended in 0.1 M potassium phosphate (pH 7.2), and stored at -80°C. It is subsequently referred to as the crude yeast microsomal fraction.

Lipid substrates. Radio-labeled ricinoleic (12-hydroxy-9-octadecenoic) and  
20    vernolic (12,13-epoxy-9-octadecenoic) acids were synthesized enzymatically from [1-<sup>14</sup>C]oleic acid and [1-<sup>14</sup>C]linoleic acid, respectively, by incubation with microsomal preparations from seeds of *Ricinus communis* and *Crepis palaestina*, respectively (12). The synthesis of phosphatidylcholines (PC) or phosphatidylethanolamines (PE) with <sup>14</sup>C-labeled acyl groups in the *sn*-2  
25    position was performed using either enzymatic (13), or synthetic (14) acylation of [<sup>14</sup>C]oleic, [<sup>14</sup>C]ricinoleic, or [<sup>14</sup>C]vernolic acid. Dioleoyl-PC that was labeled in the *sn*-1 position was synthesized from *sn*-1-[<sup>14</sup>C]oleoyl-lyso-PC and unlabeled oleic acid as described in (14). *Sn*-1-oleoyl-*sn*-2-[<sup>14</sup>C]ricinoleoyl-DAG was synthesized from PC by the action of phospholipase C type XI from *B.*  
30    *Cereus* (Sigma Chemical Co.) as described in (15). Monovernoloyl- and divernoloyl-DAG were synthesized from TAG extracted from seeds of *Euphorbia lagascae*, using the TAG-lipase (*Rizhopus arrizus*, Sigma Chemical

Co.) as previously described (16). Monoricinoleoyl-TAG was synthesized according to the same method using TAG extracted from Castor bean.

Lipid analysis. Total lipid composition of yeast were determined from cells  
5 harvested from a 40 ml liquid culture, broken in a glass-bead shaker and  
extracted into chloroform as described by Bligh and Dyer (17), and then  
separated by thin layer chromatography in hexane/diethylether/acetic acid  
(80:20:1) using pre-coated silica gel 60 plates (Merck). The lipid areas were  
located by brief exposure to I<sub>2</sub> vapors and identified by means of appropriate  
10 standards. Polar lipids, sterol-esters and triacylglycerols, as well as the  
remaining minor lipid classes, referred to as other lipids, were excised from the  
plates. Fatty acid methylesters were prepared by heating the dry excised  
material at 85 °C for 60 min in 2% (v/v) sulfuric acid in dry methanol. The  
methyl esters were extracted with hexane and analyzed by GLC through a 50 m  
15 x 0.32 mm CP-Wax58-CB fused-silica column (Chrompack), with  
methylheptadecanoic acid as an internal standard. The fatty acid content of  
each fraction was quantified and used to calculate the relative amount of each  
lipid class. In order to determine the total lipid content, 3 ml aliquots from yeast  
cultures were harvested by centrifugation and the resulting pellets were washed  
20 with distilled water and lyophilized. The weight of the dried cells was  
determined and the fatty acid content was quantified by GLC-analyses after  
conversion to methylesters as described above. The lipid content was then  
calculated as nmol fatty acid (FA) per mg dry weight yeast.

25 Enzyme assays. Aliquots of crude microsomal fractions (corresponding to  
10 nmol of microsomal PC) from developing plant seeds or yeast cells were  
lyophilized over night. <sup>14</sup>C-Labeled substrate lipids dissolved in benzene were  
then added to the dried microsomes. The benzene was evaporated under a  
stream of N<sub>2</sub>, leaving the lipids in direct contact with the membranes, and 0.1  
30 ml of 50 mM potassium phosphate (pH 7.2) was added. The suspension was  
thoroughly mixed and incubated at 30°C for the time period indicated, up to 90  
min. Lipids were extracted from the reaction mixture using chloroform and  
separated by thin layer chromatography in hexane/diethylether/acetic acid



(35:70:1.5) using silica gel 60 plates (Merck). The radioactive lipids were visualized and quantified on the plates by electronic autoradiography (Instant Imager, Packard, US).

5        Yeast cultivation. Yeast cells were grown at 28°C on a rotatory shaker in liquid YPD medium (1% yeast extract, 2% peptone, 2% glucose), synthetic medium (18) containing 2% (v/v) glycerol and 2% (v/v) ethanol, or minimal medium (19) containing 16 g/l of glycerol.

10      The instant invention is further characterized by the following examples which are not limiting:

Acyl-CoA-independent synthesis of TAG by oil seed microsomes. A large number of unusual fatty acids can be found in oil seeds (20). Many of these  
15      fatty acids, such as ricinoleic (21) and vernolic acids (22), are synthesized using phosphatidylcholin (PC) with oleoyl or linoleoyl groups esterified to the *sn*-2 position, respectively, as the immediate precursor. However, even though PC can be a substrate for unusual fatty acid synthesis and is the major membrane lipids in seeds, unusual fatty acids are rarely found in the  
20      membranes. Instead, they are mainly incorporated into the TAG. A mechanism for efficient and selective transfer of these unusual acyl groups from PC into TAG must therefore exist in oil seeds that accumulate such unusual fatty acids. This transfer reaction was biochemically characterized in seeds from castor bean (*Ricinus communis*) and *Crepis palaestina*, plants which accumulate high  
25      levels of ricinoleic and vernolic acid, respectively, and sunflower (*Helianthus annuus*), a plant which has only common fatty acids in its seed oil. Crude microsomal fractions from developing seeds were incubated with PC having <sup>14</sup>C-labeled oleoyl, ricinoleoyl or vernoloyl groups at the *sn*-2 position. After the incubation, lipids were extracted and analyzed by thin layer chromatography.  
30      We found that the amount of radioactivity that was incorporated into the neutral lipid fraction increased linearly over a period of 4 hours (data not shown). The distribution of [<sup>14</sup>C]acyl groups within the neutral lipid fraction was analyzed after 80 min (Fig. 1). Interestingly the amount and distribution of radioactivity

between different neutral lipids were strongly dependent both on the plant species and on the type of [ $^{14}\text{C}$ ]acyl chain. Thus, sunflower microsomes incorporated most of the label into DAG, regardless of the type of [ $^{14}\text{C}$ ]acyl group. In contrast, *R. communis* microsomes preferentially incorporated

5 [ $^{14}\text{C}$ ]ricinoleoyl and [ $^{14}\text{C}$ ]vernoloyl groups into TAG, while [ $^{14}\text{C}$ ]oleyl groups mostly were found in DAG. *C. palaestina* microsomes, finally, incorporated only [ $^{14}\text{C}$ ]vernoloyl groups into TAG, with [ $^{14}\text{C}$ ]ricinoleoyl groups being found mostly as free fatty acids, and [ $^{14}\text{C}$ ]oleyl groups in DAG. This shows that the high *in vivo* levels of ricinoleic acid and vernolic acid in the TAG pool of *R. communis*

10 and *C. palaestina*, respectively, can be explained by an efficient and selective transfer of the corresponding acyl groups from PC to TAG in these organisms.

The in-vitro synthesis of triacylglycerols in microsomal preparations of developing castor bean is summarized in table 1.

15

PDAT: a novel enzyme that catalyzes acyl-CoA independent synthesis of TAG. It was investigated if DAG could serve both as an acyl donor as well as an acyl acceptor in the reactions catalyzed by the oil seed microsomes. Therefore, unlabeled divernoloyl-DAG was incubated with either *sn*-1-oleoyl-

20 *sn*-2-[ $^{14}\text{C}$ ]ricinoleoyl-DAG or *sn*-1-oleoyl-*sn*-2-[ $^{14}\text{C}$ ]ricinoleoyl-PC in the presence of *R. communis* microsomes. The synthesis of TAG molecules containing both [ $^{14}\text{C}$ ]ricinoleoyl and vernoloyl groups was 5 fold higher when [ $^{14}\text{C}$ ]ricinoleoyl-PC served as acyl donor as compared to [ $^{14}\text{C}$ ]ricinoleoyl-DAG (fig.1B). These data strongly suggests that PC is the immediate acyl donor and

25 DAG the acyl acceptor in the acyl-CoA-independent formation of TAG by oil seed microsomes. Therefore, this reaction is catalyzed by a new enzyme which we call phospholipid : diacylglycerol acyltransferase (PDAT).

PDAT activity in yeast microsomes. Wild type yeast cells were cultivated

30 under conditions where TAG synthesis is induced. Microsomal membranes were prepared from these cells and incubated with *sn*-2-[ $^{14}\text{C}$ ]ricinoleoyl-PC and DAG and the  $^{14}\text{C}$ -labeled products formed were analyzed. The PC-derived [ $^{14}\text{C}$ ]ricinoleoyl groups within the neutral lipid fraction mainly were found in free

fatty acids or TAG, and also that the amount of TAG synthesized was dependent on the amount of DAG that was added to the reaction (Fig.2). The *in vitro* synthesis of TAG containing both ricinoleoyl and vernoloyl groups, a TAG species not present *in vivo*, from exogenous added *sn*-2-[<sup>14</sup>C]ricinoleoyl-PC and  
5 unlabelled vernoloyl-DAG (Fig. 2, lane 3) clearly demonstrates the existence of an acyl-CoA-independent synthesis of TAG involving PC and DAG as substrates in yeast microsomal membranes. Consequently, TAG synthesis in yeast can be catalyzed by an enzyme similar to the PDAT found in plants.

10 *The PDAT encoding gene in yeast.*

A gene in the yeast genome (YNR008w) is known, but nothing is known about the function of YNR008w, except that the gene is not essential for growth under normal circumstances. Microsomal membranes were prepared from the yeast strain FVKT004-04C(AL) (8) in which this gene with unknown function had been  
15 disrupted. PDAT activity in the microsomes were assayed using PC with radiolabelled fatty acids at the *sn*-2 position. The activity was found to be completely absent in the disruption strain (Fig. 2 lane 4). Significantly, the activity could be partially restored by the presence of YNR008w on the single copy plasmid pUS1 (Fig. 2 lane 5). Moreover, acyl groups of  
20 phosphatidylethanolamine (PE) were efficiently incorporated into TAG by microsomes from the wild type strain whereas no incorporation occurred from this substrate in the mutant strain (data not shown). This shows that YNR008w encodes a yeast PDAT which catalyzes the transfer of an acyl group from the *sn*-2 position of phospholipids to DAG, thus forming TAG. It should be noted  
25 that no cholesterol esters were formed from radioactive PC even in incubations with added ergosterols, nor were the amount of radioactive free fatty acids formed from PC affected by disruption of the YNR008w gene (data not shown). This demonstrates that yeast PDAT do not have cholesterol ester synthesising or phospholipase activities.

30

*Increased TAG content in yeast cells that overexpress PDAT.* The effect of overexpressing the PDAT-encoding gene was studied by transforming a wild type yeast strain with the pUS4 plasmid in which the gene is expressed from

the galactose-induced *GAL1:TPK2* promoter. Cells containing the empty expression vector were used as a control. The cells were grown in synthetic glycerol-ethanol medium, and expression of the gene was induced after either 2 hours (early log phase) or 25 hours (stationary phase) by the addition of galactose. The cells were then incubated for another 21 hours, after which they were harvested and assays were performed. We found that overexpression of PDAT had no significant effect on the growth rate as determined by the optical density. However, the total lipid content, measured as  $\mu\text{mol}$  fatty acids per mg yeast dry weight, was 47% (log phase) or 29% (stationary phase) higher in the PDAT overexpressing strain than in the control. Furthermore, the polar lipid and sterolester content was unaffected by overexpression of PDAT. Instead, the elevated lipid content in these cells is entirely due to an increased TAG content (Fig. 3A,B). Thus, the amount of TAG was increased by 2-fold in PDAT overexpressing early log phase cells and by 40% in stationary phase cells. It is interesting to note that a significant increase in the TAG content was achieved by overexpressing PDAT even under conditions (*i.e.* in stationary phase) where DAGAT is induced and thus contributes significantly to TAG synthesis. *In vitro* PDAT activity assayed in microsomes from the PDAT overexpressing strain was 7-fold higher than in the control strain, a finding which is consistent with the increased levels of TAG that we observed *in vivo* (Fig. 3C). These results clearly demonstrate the potential use of the PDAT gene in increasing the oil content in transgenic organisms.

Substrate specificity of yeast PDAT. The substrate specificity of yeast PDAT was analyzed using microsomes prepared from the PDAT overexpressing strain (see Fig. 4). The rate of TAG synthesis, under conditions given in figure 4 with di-oleoyl-PC as the acyl-donor, was 0.15 nmol per min and mg protein. With both oleoyl groups of PC labeled it was possible, under the given assay conditions, to detect the transfer of 11 pmol/min of [ $^{14}\text{C}$ ]oleoyl chain into TAG and the formation of 15 pmol/min of lyso-PC. In microsomes from the PDAT-deficient strain, no TAG at all and only trace amounts of lyso-PC was detected, strongly suggesting that yeast PDAT catalyses the formation of equimolar amounts of TAG and lyso-PC when supplied with PC and DAG as

substrates. The fact that somewhat more lyso-PC than TAG is formed can be explained by the presence of a phospholipase in yeast microsomes, which produces lyso-PC and unesterified fatty acids from PC.

5       The specificity of yeast PDAT for different acyl group positions was investigated by incubating the microsomes with di-oleoyl-PC carrying a [ $^{14}\text{C}$ ]acyl group either at the *sn*-1 position (Fig. 4A bar 2) or the *sn*-2 position (Fig. 4A bar 3). We found that the major  $^{14}\text{C}$ -labeled product formed in the former case was lyso-PC, and in the latter case TAG. We conclude that yeast  
10   PDAT has a specificity for the transfer of acyl groups from the *sn*-2 position of the phospholipid to DAG, thus forming *sn*-1-lyso-PC and TAG. Under the given assay conditions, trace amounts of  $^{14}\text{C}$ -labelled DAG is formed from the *sn*-1 labeled PC by the reversible action of a CDP-choline : choline phosphotransferase. This labeled DAG can then be further converted into TAG  
15   by the PDAT activity. It is therefore not possible to distinguish whether the minor amounts of labeled TAG that is formed in the presence of di-oleoyl-PC carrying a [ $^{14}\text{C}$ ]acyl group in the *sn*-1 position, is synthesized directly from the *sn*-1-labeled PC by a PDAT that also can act on the *sn*-1 position, or if it is first converted to *sn*-1-labeled DAG and then acylated by a PDAT with strict  
20   selectivity for the transfer of acyl groups at the *sn*-2 position of PC. Taken together, this shows that the PDAT encoded by YNR008w catalyses an acyl transfer from the *sn*-2 position of PC to DAG, thus causing the formation of TAG and lyso-PC.

25       The substrate specificity of yeast PDAT was further analyzed with respect to the headgroup of the acyl donor, the acyl group transferred and the acyl chains of the acceptor DAG molecule. The two major membrane lipids of *S. cerevisiae* are PC and PE, and as shown in Fig. 4B (bars 1 and 2), dioleoyl-PE is nearly 4-fold more efficient than dioleoyl-PC as acyl donor in the PDAT-catalyzed reaction. Moreover, the rate of acyl transfer is strongly dependent on  
30   the type of acyl group that is transferred. Thus, a ricinoleoyl group at the *sn*-2 position of PC is 2.5 times more efficiently transferred into TAG than an oleoyl group in the same position (Fig. 4B bars 1 and 3). In contrast, yeast PDAT has

no preference for the transfer of vernoloyl groups over oleoyl groups (Fig. 4B bars 1 and 4). The acyl chain of the acceptor DAG molecule also affects the efficiency of the reaction. Thus, DAG with a ricinoleoyl or a vernoloyl group is a more efficient acyl acceptor than dioleoyl-DAG (Fig. 4B bars 1, 5 and 6). Taken  
5 together, these results clearly show that the efficiency of the PDAT-catalyzed acyl transfer is strongly dependent on the properties of the substrate lipids.

PDAT genes. Nucleotide and amino acid sequences of several PDAT genes are given as SEQ ID No. 1 through 15. Further provisional and/or partial  
10 sequences are given as SEQ ID NO 16 through 20 and 21 through 31, respectively. One of the Arabidopsis genomic sequences (SEQ ID NO. 4) identified an Arabidopsis EST cDNA clone; T04806. This cDNA clone was fully characterised and the nucleotide sequence is given as SEQ ID NO. 5. Based  
15 on the sequence homology of the T04806 cDNA and the *Arabidopsis thaliana* genomic DNA sequence (SEQ ID NO 4) it is apparent that an additional A is present at position 417 in the cDNA clone (data not shown). Excluding this nucleotide would give the amino acid sequence depicted in SEQ ID NO. 12.

Increased TAG content in seeds of Arabidopsis thaliana that express the  
20 yeast PDAT. For the expression of the yeast PDAT gene in *Arabidopsis thaliana* an EcoRI fragment from the pBluescript-PDAT was cloned together with napin promotor (25) into the vector pGPTV-KAN (26). A plasmid (pGNapPDAT) having the yeast PDAT gene in the correct orientation was identified and transformed into *Agrobacterium tumefaciens*. These bacteria  
25 were used to transform *Arabidopsis thaliana* columbia (C-24) plants using the root transformation method (27). Plants transformed with an empty vector were used as controls.

First generation seeds (T1) were harvested and germinated on kanamycin containing medium. Second generation seeds (T2) were pooled from individual  
30 plants and their fatty acid contents analysed by quantification of their methyl esters by gas liquid chromatography after methylation of the seeds with 2% sulphuric acid in methanol at 85 °C for 1,5 hours. Quantification was done with heptadecanoic acid methyl esters as internal standard.

From the transformation with pGNapPDAT one T1 plant (26-14) gave raise to seven T2 plants of which 3 plants yielded seeds with statistically (in a mean difference two-sided test) higher oil content than seeds from T2 plants generated from T1 plant 32-4 transformed with an empty vector (table 2).

## References cited in the description:

1. Bell, R. M. & Coleman, R. A. (1980) *Annu. Rev. Biochem.* **49**, 459-487.
2. Stymne, S. & Stobart, K. (1987) in *The biochemistry of plants: a comprehensive treatise*, Vol. 9, ed. Stumpf, P. K. (Academic Press, New York), pp. 175-214.
3. Cases, S. et al. (1998) *Proc. Natl. Acad. Sci. U S A* **95**, 13018-13023.
4. Hobbs, D. H., Lu, C. & Hills, M. J. (1999) *FEBS Lett.* **452**, 145-9
5. Zou, J., Wei, Y., Jako, C., Kumar, A., Selvaraj, G. & Taylor, D. C. (1999) *Plant J.* **19**, 645-653.
6. Lardizabal, K., Hawkins, D., Mai, J., & Wagner, N. (1999) Abstract presented at the Biochem. Mol. Plant Fatty Acids Glycerolipids Symposium, South Lake Tahoe, USA.
7. Thomas, B. J. & Rothstein, R. (1989) *Cell* **56**, 619-630.
8. Entian, K.-D. & Kötter, P. (1998) *Meth. Microbiol.* **26**, 431-449.
9. Kern, L., de Montigny, J., Jund, R. & Lacroute, F. (1990) *Gene* **88**, 149-157.
10. Ronne, H., Carlberg, M., Hu, G.-Z. & Nehlin, J. O. (1991) *Mol. Cell. Biol.* **11**, 4876-4884.
11. Stobart, K. & Stymne, S. (1990) in *Method in Plant Biochemistry*, vol 4, eds. Harwood, J. L. & Bowyer, J. R. (Academic press, London), pp. 19-46.
12. Bafor, M., Smith, M. A., Jonsson, L., Stobrt, A. K. & Stymne, S. (1991) *Biochem. J.* **280**, 507-514.
13. Banas, A., Johansson, I. & Stymne, S. (1992) *Plant Science* **84**, 137-144.
14. Kanda, P. & Wells, M. A. (1981) *J. Lipid. Res.* **22**, 877-879.
15. Ståhl, U., Ek, B. & Stymne, S. (1998) *Plant Physiol.* **117**, 197-205.
16. Stobart, K., Mancha, M. & Lenman M, Dahqvist, A. & Stymne, S. (1997) *Planta* **203**, 58-66.
17. Bligh, E. G. & Dyer, W. J. (1959) *Can. J. Biochem. Physiol.* **37**, 911-917.
18. Sherman, F., Fink, G. R. & Hicks, J. B. (1986) in *Laboratory Course Manual for Methods in Yeast Genetics* (Cold Spring Harbor Laboratory)
19. Meesters, P. A. E. P., Huijberts, G. N. M. and Eggink, G. (1996) *Appl. Microbiol. Biotechnol.* **45**, 575-579.
20. van de Loo, F. J., Fox, B. G. & Sommerville, C. (1993), in *Lipid metabolism in plants*, ed. Moore, T. S. (CRC Press, Inc.), pp. 91-126.
21. van de Loo, F. J., Broun, P., Turner, S. & Sommerville, S. (1995) *Proc. Natl. Acad. Sci. U S A* **95**, 6743-6747.
22. Lee, M., Lenman, M., Banas, A., Bafor, M., Singh, S., Schweizer, M., Nilsson, R., Liljenberg, C., Dahqvist, A., Gummeson, P.-O., Sjö Dahl, S.,



- Green, A., and Stymne, S. (1998) *Science* **280**, 915-918.
23. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997) *Nucl. Acids Res.* **24**, 4876-4882.
24. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4**, 406-425.
- 5 25. Stålberg, K., Ellerström, M., Josefsson, L., & Rask, L. (1993) *Plant Mol. Biol.* **23**, 671
26. Becker, D., Kemper, E., Schell, J., Masterson, R. (1992) *Plant Mol. Biol.* **20**, 1195
27. D. Valvekens, M. Van Montagu, and Van Lusbettens (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5536
- 10

## Description of Figures

FIG. 1.

Metabolism of  $^{14}\text{C}$ -labeled PC into the neutral lipid fraction by plant  
5 microsomes. (A) Microsomes from developing seeds of sunflower, *R.*  
*communis* and *C. palaestina* were incubated for 80 min at 30°C with PC (8  
nmol) having oleic acid in its *sn*-1 position, and either  $^{14}\text{C}$ -labeled oleic,  
ricinoleic or vernolic acid in its *sn*-2 position. Radioactivity incorporated in TAG  
(open bars), DAG (solid bars), and unsterified fatty acids (hatched bars) was  
10 quantified using thin layer chromatography followed by electronic  
autoradiography, and is shown as percentage of added labeled substrate. (B)  
Synthesis *in vitro* of TAG carrying two vernoloyl and one [ $^{14}\text{C}$ ]ricinoleoyl group  
by microsomes from *R. communis*. The substrates added were unlabeled  
divernoloyl-DAG (5 nmol), together with either *sn*-1-oleoyl-*sn*-2-[ $^{14}\text{C}$ ]ricinoleoyl-  
15 DAG (0.4 nmol, 7700 dpm/nmol) or *sn*-1-oleoyl-*sn*-2-[ $^{14}\text{C}$ ]ricinoleoyl-PC (0.4  
nmol, 7700 dpm/nmol). The microsomes were incubated with the substrates for  
30 min at 30°C, after which samples were removed for lipid analysis as  
described in the section „general methods“. The data shown are the average of  
two experiments.

20

FIG. 2.

PDAT activity in yeast microsomes, as visualized by autoradiogram of neutral  
lipid products separated on TLC. Microsomal membranes (10 nmol of PC) from  
the wild type yeast strain FY1679 (lanes 1-3), a congenic yeast strain  
25 (FVKT004-04C(AL)) that is disrupted for YNR008w (lane 4) or the same  
disruption strain transformed with the plasmid pUS1, containing the YNR008w  
gene behind its native promotor (lane 5), were assayed for PDAT activity. As  
substrates, we used 2 nmol *sn*-1-oleoyl-*sn*-2-[ $^{14}\text{C}$ ]ricinoleoyl-PC together with  
either 5 nmol of dioleoyl-DAG (lanes 2, 4 and 5) or *rac*-oleoyl-vernoleoyl-DAG  
30 (lane 3). The enzymatic assay and lipid analysis was performed as described in  
Materials and Methods. The cells were precultured for 20 h in liquid YPD  
medium, harvested and re-suspended in an equal volume of minimal medium  
(19) containing 16 g/l glycerol. The cells were then grown for an additional 24 h

prior to being harvested. Selection for the plasmid was maintained by growing the transformed cells in synthetic medium lacking uracil (18). Abbreviations: 1-OH-TAG, monoricinoleoyl-TAG; 1-OH-1-ep-TAG, monoricinoleoyl-monovernoloyl-TAG; OH-FA, unesterified ricinoleic acid.

5

Fig. 3.

Lipid content (A,B) and PDAT activity (C) in PDAT overexpressing yeast cells.

The PDAT gene in the plasmid pUS4 was overexpressed from the galactose-induced *GAL1-TPK2* promotor in the wild type strain W303-1A (7). Its expression was induced after (A) 2 hours or (B) 25 hours of growth by the addition of 2% final concentration (w/v) of galactose. The cells were then incubated for another 22 hours before being harvested. The amount of lipids of the harvested cells was determined by GLC-analysis of its fatty acid contents and is presented as  $\mu\text{mol}$  fatty acids per mg dry weight in either TAG (open bar), polar lipids (hatched bar), sterol esters (solid bar) and other lipids (striped bar). The data shown are the mean values of results with three independent yeast cultures. (C) *In vitro* synthesis of TAG by microsomes prepared from yeast cells containing either the empty vector (vector) or the PDAT plasmid (+ PDAT). The cells were grown as in Fig. 3A. The substrate lipids dioleoyl-DAG (2.5 nmol) and *sn*-1-oleoyl-*sn*-2- $^{14}\text{C}$ -oleoyl-PC (2 nmol) were added to aliquots of microsomes (10 nmol PC), which were then incubated for 10 min at 28 °C. The amount of label incorporated into TAG was quantified by electronic autoradiography. The results shown are the mean values of two experiments.

25 FIG. 4.

Substrate specificity of yeast PDAT. The PDAT activity was assayed by incubating aliquots of lyophilized microsomes (10 nmol PC) with substrate lipids at 30°C for 10 min (panel A) or 90 min (panel B). Unlabeled DAG (2.5 nmol) was used as substrates together with different labeled phospholipids, as shown in the figure. (A) *Sn*-position specificity of yeast PDAT regarding the acyl donor substrate. Dioleoyl-DAG together with either *sn*-1- $^{14}\text{C}$ -oleoyl-*sn*-2- $^{14}\text{C}$ -oleoyl-PC (di- $^{14}\text{C}$ -PC), *sn*-1- $^{14}\text{C}$ -oleoyl-*sn*-2-oleoyl-PC (*sn*1- $^{14}\text{C}$ -PC) or *sn*-1-oleoyl-*sn*-2- $^{14}\text{C}$ -oleoyl-PC (*sn*2- $^{14}\text{C}$ -PC). (B) Specificity of yeast PDAT regarding

30

phospholipid headgroup and of the acyl composition of the phospholipid as well as of the diacylglycerol. Dioleoyl-DAG together with either *sn*-1-oleoyl-*sn*-2-[<sup>14</sup>C]oleoyl-PC (oleoyl-PC), *sn*-1-oleoyl-*sn*-2-[<sup>14</sup>C]oleoyl-PE (oleoyl-PE), *sn*-1-oleoyl-*sn*-2-[<sup>14</sup>C]ricinoleoyl-PC (ricinoleoyl-PC) or *sn*-1-oleoyl-*sn*-2-[<sup>14</sup>C]vernoloyl-PC (vernoloyl-PC). In the experiments presented in the 2 bars to the far right, monoricinoleoyl-DAG (ricinoleoyl-DAG or mono-vernoloyl-DAG (vernoloyl-DAG) were used together with *sn*-1-oleoyl-*sn*-2-[<sup>14</sup>C]-oleoyl-PC. The label that was incorporated into TAG (solid bars) and lyso-PC (LPC, open bars) was quantified by electronic autoradiography. The results shown are the mean values of two experiments. The microsomes used were from W303-1A cells overexpressing the PDAT gene from the *GAL1-TPK2* promotor, as described in Fig. 3. The expression was induced at early stationary phase and the cells were harvested after an additional 24 h.

15

## TAB.1:

In vitro synthesis of triacylglycerols in microsomal preparations of developing castor bean. Aliquots of microsomes (20 nmol PC) were lyophilised and substrate lipids were added in benzene solution: (A) 0.4 nmol [<sup>14</sup>C]-DAG (7760 dpm/nmol) and where indicated 1.6 nmol unlabelled DAG; (B) 0.4 nmol [<sup>14</sup>C]-DAG (7760 dpm/nmol) and 5 nmol unlabelled di-ricinoleoyl-PC and (C) 0.25 nmol [<sup>14</sup>C]-PC (4000 dpm/nmol) and 5 nmol unlabelled DAG. The benzene was evaporated by N<sub>2</sub> and 0.1 ml of 50 mM potassium phosphate was added, thoroughly mixed and incubated at 30 °C for (A) 20 min.; (B) and (C) 30 min.. Assays were terminated by extraction of the lipids in chloroform. The lipids were then separated by thin layer chromatography on silica gel 60 plates (Merck; Darmstadt, Germany) in hexan/diethylether/acetic 35:70:1.5. The radioactive lipids were visualised and the radioactivity quantified on the plate by electronic autoradiography (Instant Imager, Packard, US). Results are presented as mean values of two experiments.

30

Radioactivity in different triacylglycerols (TAG) species formed. Abbreviations used: 1-OH-, mono-ricinoleoyl-; 2-OH-, di-ricinoleoyl-; 3-OH-, triricinoleoyl-; 1-

OH-1-ver-, mono-ricinoleoyl-monovernoleoyl-; 1-OH-2-ver-, mono-ricinoleoyl-divernoleoyl-. Radiolabelled DAG and PC were prepared enzymatically. The radiolabelled ricinoleoyl group is attached at the sn-2-position of the lipid and unlabelled oleoyl group at the sn-1-position. Unlabelled DAG with vernoleoyl- or  
5 ricinoleoyl chains were prepared by the action of TAG lipase (6) on oil of *Euphorbia lagascae* or Castor bean, respectively. Synthetic di-ricinoleoyl-PC was kindly provided from Metapontum Agribios (Italy).

10 TAB.2:

Total fatty acids per mg of T2 seeds pooled from individual *Arabidopsis thaliana* plants transformed with yeast PDAT gene under the control of napin promotor (26-14) or transformed with empty vector (32-4).

15 \* = statistical difference between control plants and PDAT transformed plants in a mean difference two-sided test at  $\alpha = 5$ .

Description of the SEQ ID:

5 *SEQ ID NO. 1:* Genomic DNA sequence and suggested amino acid sequence of the *Saccharomyces cerevisiae* PDAT gene, YNR008w, with GenBank accession number Z71623 and Y13139, and with nucleotide ID number 1302481.

*SEQ ID NO. 2:* The amino acid sequence of the suggested open reading frame YNR008w from *Saccharomyces cerevisiae*.

10 *SEQ ID NO. 3:* Genomic DNA sequence of the *Schizosaccharomyces pombe* gene SPBC776.14.

*SEQ ID NO. 4:* Genomic DNA sequence of part of the *Arabidopsis thaliana* locus with GenBank accession number AB006704.

15

*SEQ ID NO. 5:* Nucleotide sequence of the *Arabidopsis thaliana* cDNA clone with GenBank accession number T04806, and nucleotide ID number 315966.

20 *SEQ ID NO. 6:* Predicted amino acid sequence of the *Arabidopsis thaliana* cDNA clone with GenBank accession number T04806.

*SEQ ID NO. 7:* Nucleotide and amino acid sequence of the *Zea mays* EST clone with GenBank accession number AI491339, and nucleotide ID number 4388167.

25 *SEQ ID NO. 8:* Predicted amino acid sequence of the *Zea mays* EST clone with GenBank accession number AI491339, and nucleotide ID number 4388167.

30 *SEQ ID NO. 9:* DNA sequence of part of the *Neurospora crassa* EST clone W07G1, with GenBank accession number AI398644, and nucleotide ID number 4241729.

*SEQ ID NO. 10:* Genomic DNA sequence of part of the *Arabidopsis thaliana* locus

with GenBank accession number AC004557.

*SEQ ID NO. 11:* Genomic DNA sequence of part of the *Arabidopsis thaliana* locus with GenBank accession number AC003027.

5

*SEQ ID NO. 12:* DNA sequence of part of the *Lycopersicon esculentum* cDNA clone with GenBank accession number AI486635.

*SEQ ID NO. 13:* Amino acid sequence of the *Schizosaccharomyces pombe* putative open reading frame CAA22887 of the *Schizosaccharomyces pombe* gene SPBC776.14.

10

*SEQ ID NO. 14:* Amino acid sequence of the *Arabidopsis thaliana* putative open reading frame AAC80628 derived from the *Arabidopsis thaliana* locus with GenBank accession number AC004557.

15

*SEQ ID NO 15:* Amino acid sequence of the *Arabidopsis thaliana* putative open reading frame AAD10668 derived from the *Arabidopsis thaliana* locus with GenBank accession number AC003027.

20

Further provisional and/or partial sequences are defined through the following SEQ IDs:

*SEQ ID NO. 16:* The amino acid sequence of the yeast ORF YNR008w from *Saccharomyces cerevisiae*.

25

*SEQ ID NO. 17:* Amino acid sequence of the region of the *Arabidopsis thaliana* genomic sequence (AC004557).

*SEQ ID NO. 18:* Amino acid sequence of the region of the *Arabidopsis thaliana* genomic sequence (AB006704).

30

SEQ ID NO. 19: The corresponding genomic DNA sequence and amino acid sequence of the yeast ORF YNROO8w from *Saccharomyces cerevisiae*.

SEQ ID NO. 20: The amino acid sequence of the yeast ORF YNROO8w from  
5 *Saccharomyces cerevisiae* derived from the corresponding genomic DNA sequence.

SEQ ID NO. 21: Genomic DNA sequence of the *Saccharomyces cerevisiae*  
10 PDAT gene, YNR008w, genebank nucleotide ID number 1302481, and the suggested YNR008w amino acid sequence.

SEQ ID NO. 22: The suggested amino acid sequence of the yeast gene  
YNR008w from *Saccharomyces cerevisiae*.  
15

SEQ ID NO. 23: Genomic DNA sequence of the *Schizosaccharomyces pombe*  
gene SPBC776.14.

SEQ ID NO. 24: Genomic DNA sequence of part of the *Arabidopsis thaliana*  
20 locus with genebank accession number AB006704.

SEQ ID NO. 25: Nucleotide sequence and the corresponding amino acid  
sequence of the *Arabidopsis thaliana* EST-clone with genebank accession  
number T04806, and ID number 315966.  
25

SEQ ID NO. 26: Nucleotide and amino acid sequence of the *Zea mays* cDNA  
clone with genebank ID number 4388167.

SEQ ID NO. 27: Amino acid sequence of the *Zea mays* cDNA clone with  
30 genebank ID number 4388167.

SEQ ID NO. 28: DNA sequence of part of the *Neurospora crassa* cDNA clone  
WO7G1, ID number 4241729.



*SEQ ID NO. 29:* Genomic DNA sequence of part of the *Arabidopsis thaliana* locus with genebank accession number AC004557.

- 5 *SEQ ID NO. 30:* Genomic DNA sequence of part of the *Arabidopsis thaliana* locus with genebank accession number AC003027.

*SEQ ID NO. 31:* DNA sequence of part of the *Lycopersicon esculentum* cDNA clone with genebank accession number AI486635.

**Claims**

1. An enzyme catalysing in an acyl-CoA-independent reaction the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway  
5 for the production of triacylglycerol.
2. An enzyme according to claim 1, comprising an amino acid sequence as set forth in SEQ ID No. 2 or a functional fragment, derivate, allele, homolog or isoenzyme thereof.  
10
3. An enzyme according to claims 1 or 2 designated as phospholipid:diacylglycerol acyltransferase (PDAT).
4. An enzyme according to claims 1 to 3, comprising an amino acid sequence  
15 as set forth in SEQ ID No. 16, 20 or 22 or a functional fragment, derivate, allele, homolog or isoenzyme thereof.
5. An enzyme according to claims 1 to 4, comprising an amino acid sequence selected from the group consisting of sequences as set forth in SEQ ID No.  
20 6, 8, 13, 14, 15, 17, 18, 25 or 27 or a functional fragment, derivate, allele, homolog or isoenzyme thereof.
6. An enzyme according to claims 1 to 5, comprising an amino acid sequence encoded through a nucleotide sequence, a portion, derivate, allele or  
25 homolog thereof selected from the group consisting of sequences as set forth in SEQ ID No. 1, 3, 4, 5, 7, 9, 10, 11, 12, 19, 21, 23, 24, 25, 26, 28, 29, 30 or 31 or a functional fragment, derivate, allele, homolog or isoenzyme of the enzyme encoding amino acid sequence.
- 30 7. A nucleotide sequence encoding an enzyme catalysing in an acyl-CoA-independent reaction the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol.

8. A nucleotide sequence according to claim 7 encoding an enzyme designated as phospholipid:diacylglycerol acyltransferase (PDAT).
- 5 9. A nucleotide sequence according to claims 7 or 8, selected from the group consisting of sequences as set forth in SEQ ID No. 1, 3, 4, 10, 11, 19, 21, 23, 24, 29 or 30 or a portion, derivate, allele or homolog thereof.
- 10 10. A partial nucleotide sequence corresponding to a fulllength nucleotide sequence according to claims 7 to 9, selected from the group consisting of sequences as set forth in SEQ ID No. 5, 7, 9, 12, 25, 26, 28 or 31 or a portion, derivate, allele or homolog thereof.
- 15 11. A nucleotide sequence according to claims 7 to 10, comprising a nucleotide sequence which is at least 40% homologous to a nucleotide sequence selected from the group consisting of those sequences set forth in SEQ ID No. 1, 3, 4, 5, 7, 9, 10, 11, 12, 19, 21, 23, 24, 25, 26, 28, 29, 30 or 31.
- 20 12. A gene construct comprising a nucleotide sequence according to claims 7 to 11 operably linked to a heterologous nucleic acid.
13. A vector comprising a nucleotide sequence according to claims 7 to 11 or a gene construct according to claim 12.
- 25 14. A vector according to claim 13, which is an expression vector.
15. A vector according to claims 13 or 14, further comprising a selectable marker gene and/or nucleotide sequences for the replication in a host cell or the integration into the genome of the host cell.
- 30 16. A transgenic cell or organism containing a nucleotide sequence according to claims 7 to 11 and/or a gene construct according to claim 12 and/or a vector according to claims 13 to 15.

17. A transgenic cell or organism according to claim 16 which is an eucaryotic cell or organism.
- 5 18. A transgenic cell or organism according to claims 16 or 17 which is a yeast cell or a plant cell or a plant.
19. A transgenic cell or organism according to claims 16 to 18 having an altered biosynthetic pathway for the production of triacylglycerol.
- 10 20. A transgenic cell or organism according to claims 16 to 19 having an altered oil content.
21. A transgenic cell or organism according to claims 16 to 20 wherein the  
15 activity of PDAT is altered.
22. A transgenic cell or organism according to claims 16 to 21 wherein the altered activity of PDAT is characterized by an alteration in gene expression, catalytic activity and/or regulation of activity of the enzyme.
- 20 23. A transgenic cell or organism according to claims 16 to 22 wherein the altered biosynthetic pathway for the production of triacylglycerol is characterized by the prevention of accumulation of undesirable fatty acids in the membrane lipids.
- 25 24. A process for the production of triacylglycerol, comprising growing a transgenic cell or organism according to claims 16 to 23 under conditions whereby the said nucleotide sequence according to claims 7 to 11 is expressed.
- 30 25. Triacylglycerols produced by a process according to claim 24.

26. Use of a nucleotide sequence according to claims 7 to 11 and/or an enzyme according to claims 1 to 6 for the production of triacylglycerol and/or triacylglycerols with uncommon fatty acids.
- 5 27. Use of a nucleotide sequence according to claims 7 to 11 and/or an enzyme according to claims 1 to 6 for the transformation of any cell or organism in order to be expressed in this cell or organism and result in an altered, preferably increased oil content of this cell or organism.

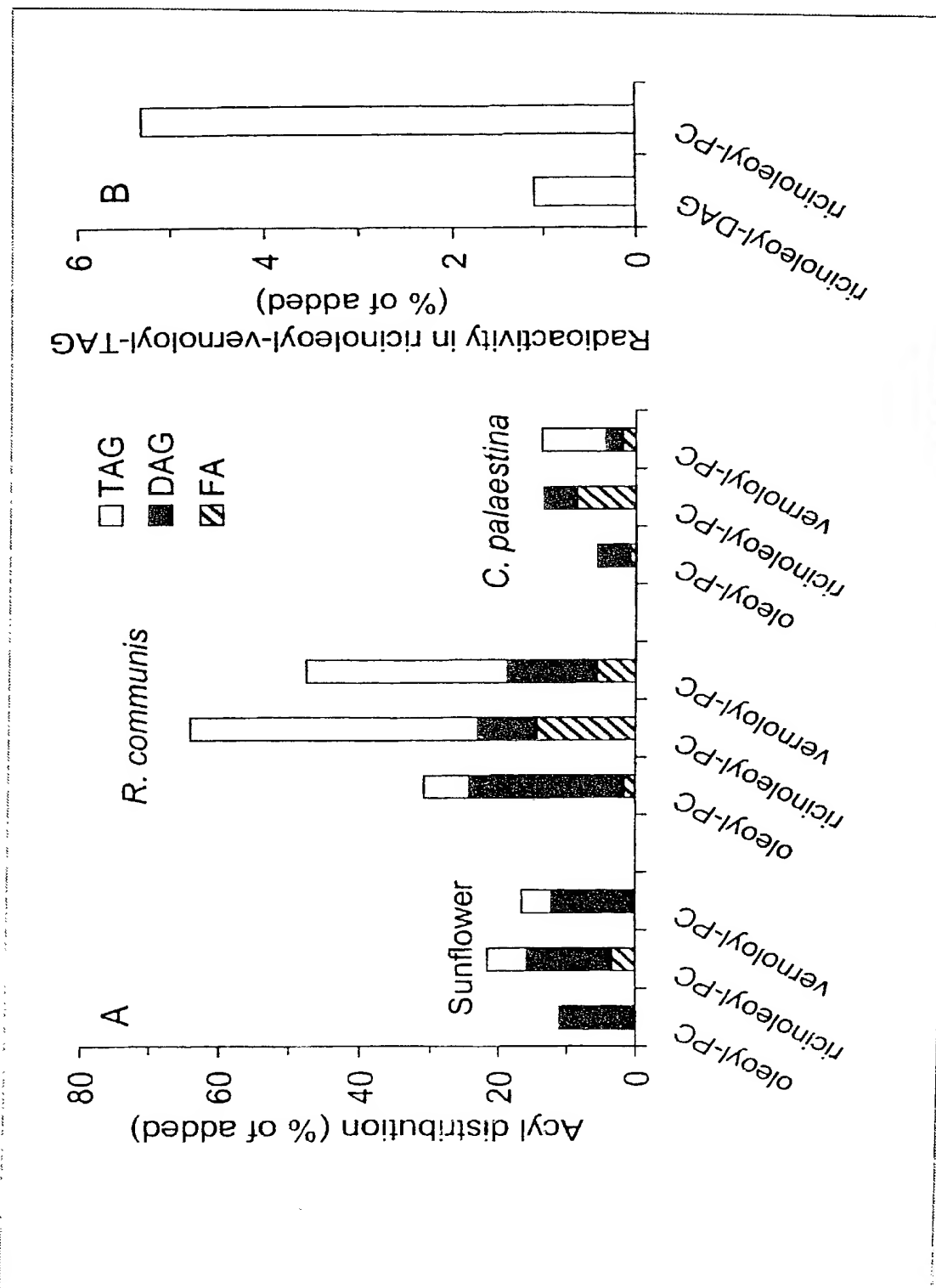
Abstract of the Disclosure

5 The present invention relates to the isolation, identification and characterization of nucleotide sequences encoding an enzyme catalysing the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol, to the said enzymes and a process for the production of triacylglycerols.

09/937779

## Figures

Fig. 1:



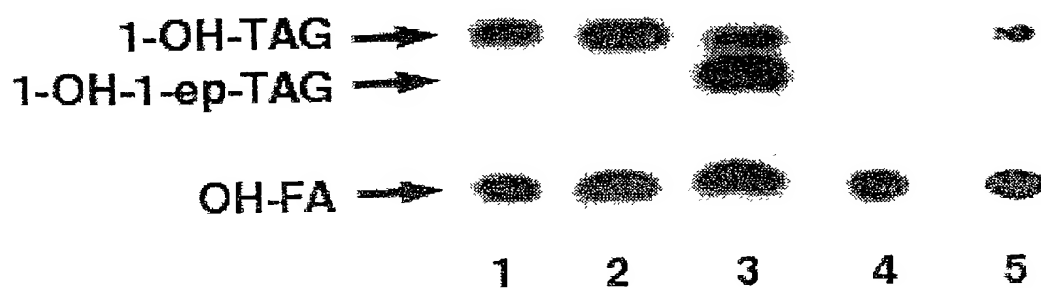
**Fig 2**



Fig 3:

09/937779

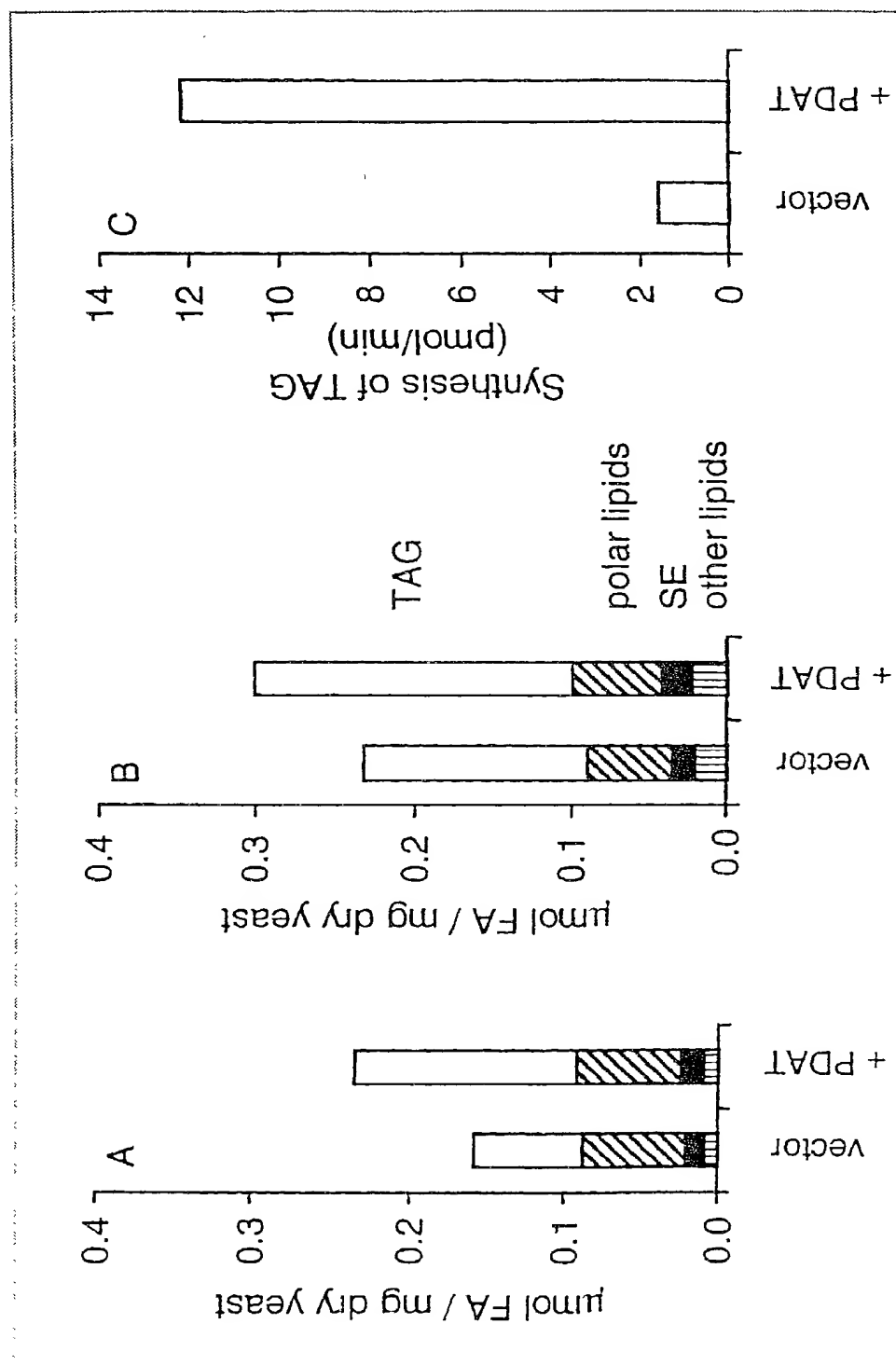
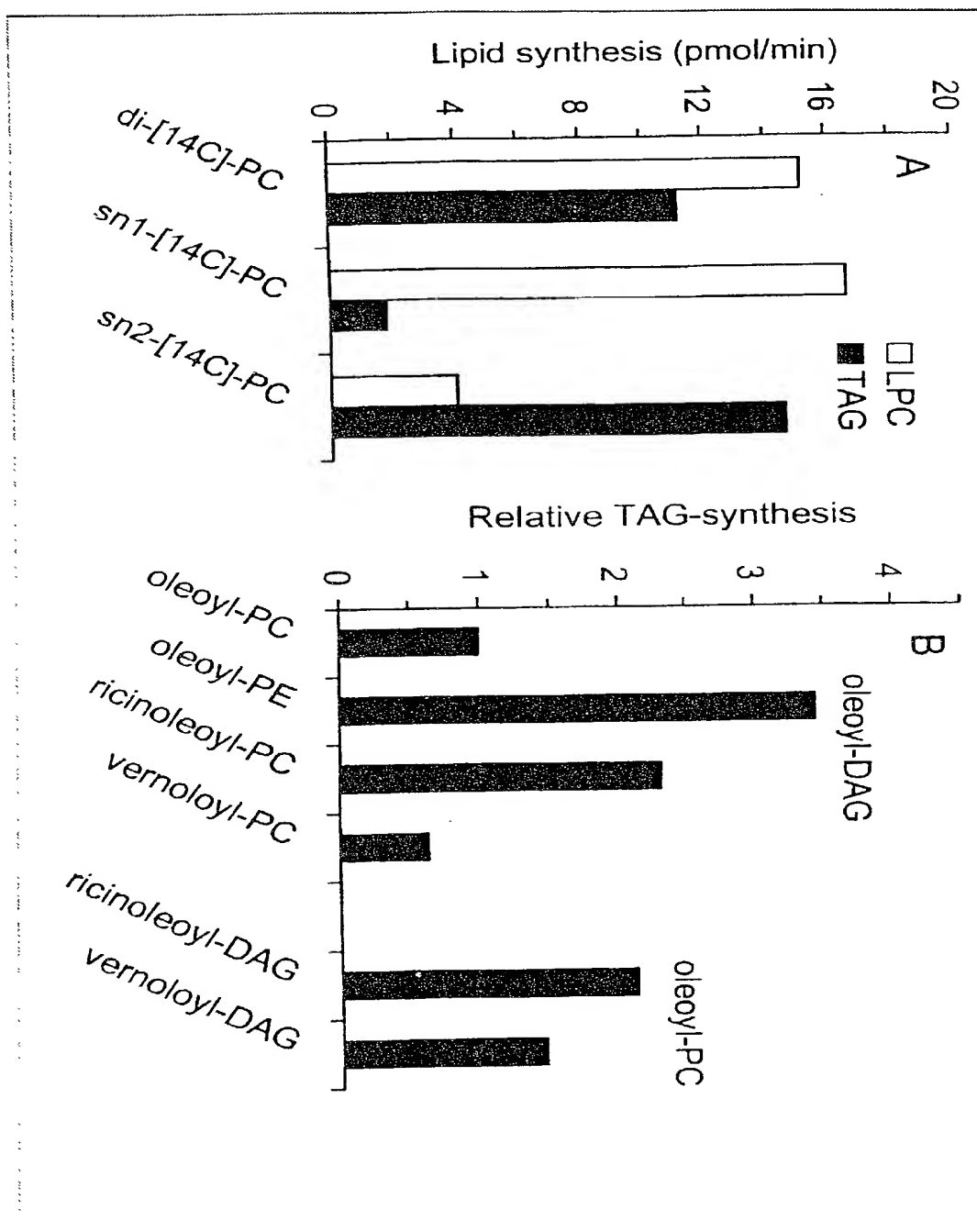


Fig.4:

09/937779



## Tables

Tab. 1:

Substrate added [ $^{14}\text{C}$ ]-lipid <sup>(2)</sup>	mol % of added [ $^{14}\text{C}$ ]-acyl group in TAG <sup>(1)</sup>					
	unlabelled lipid <sup>(2)</sup>	1-OH-TAG	2-OH-TAG	1-OH-1-ver-TAG	1-OH-2-ver-TAG	3-OH-TAG
A mono-[ $^{14}\text{C}$ ]-ricinoleoyl-DAG	mono-ricinoleoyl-DAG	2,8	12,4	-	-	-
A mono-[ $^{14}\text{C}$ ]-ricinoleoyl-DAG	mono-vernoleoyl-DAG	3,2	12,1	1,3	-	-
A mono-[ $^{14}\text{C}$ ]-ricinoleoyl-DAG	di-vernoleoyl-DAG	4	10	0,5	1,2	-
A mono-[ $^{14}\text{C}$ ]-ricinoleoyl-DAG	di-ricinoleoyl-PC	0,3	24,8	-	-	-
B mono-[ $^{14}\text{C}$ ]-ricinoleoyl-PC	none	6,8	8,0	-	-	4,7
C mono-[ $^{14}\text{C}$ ]-ricinoleoyl-PC	di-oleoyl-DAG	8,6	9,8	-	-	5,0
C mono-[ $^{14}\text{C}$ ]-ricinoleoyl-PC	mono-ricinoleoyl-DAG	5,7	16,7	-	-	1,9
C mono-[ $^{14}\text{C}$ ]-ricinoleoyl-PC	di-ricinoleoyl-DAG	4,5	9,4	-	-	9,5
C mono-[ $^{14}\text{C}$ ]-ricinoleoyl-PC	mono-vernoleoyl-DAG	6,0	11,5	10,9	0,5	7,4
C mono-[ $^{14}\text{C}$ ]-ricinoleoyl-PC	di-vernoleoyl-DAG	6,7	10,8	1,1	8,4	6,8

09/937779

09/937779

Tab. 2:

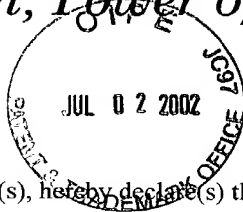
T1 plant	T2 plant number	nmol fatty acids per mg seed		standard deviation
32-4	1	1277		$\pm 11$ (n=2)
	4	1261		$\pm 63$ (n=3)
	5	1369		$\pm 17$ (n=3)
	6	1312		$\pm 53$ (n=4)
	7	1197		$\pm 54$ (n=5)
	8	1240		$\pm 78$ (n=4)
	9	1283		$\pm 54$ (n=5)
	10	1381		$\pm 35$ (n=5)
26-14	1	1444		$\pm 110$ (n=4)
	2	1617*		$\pm 109$ (n=4)
	3	1374		$\pm 37$ (n=2)
	5	1562*		$\pm 70$ (n=4)
	6	1393		$\pm 77$ (n=4)
	7	1433		$\pm 98$ (n=4)
	8	1581*		$\pm 82$ (n=4)

# Declaration, Power of Attorney and Petition



26474

PATENT & TRADEMARK OFFICE



Page 1 of 4  
0093/000003

We (I), the undersigned inventor(s), hereby declare(s) that:

My residence, post office address and citizenship are as stated below next to my name,

We (I) believe that we are (I am) the original, first, and joint (sole) inventor(s) of the subject matter which is claimed and for which a patent is sought on the invention entitled

NEW CLASS OF ENZYMES IN THE BIOSYNTHETIC PATHWAY FOR THE PRODUCTION OF  
TRIACYCLOGLYCEROL AND RECOMBINANT DNA MOLECULES ENCODING THESE ENZYMES

the specification of which

☐ is attached hereto.

☒ was filed on September 28, 2001 as

Application Serial No. 09/937,779

and amended on \_\_\_\_\_.

☒ was filed as PCT international application

Number PCT/EP/00/02701

on March 28, 2000

and was amended under PCT Article 19

on \_\_\_\_\_ (if applicable).

We (I) hereby state that we (I) have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

We (I) acknowledge the duty to disclose information known to be material to the patentability of this application as defined in Section 1.56 of Title 37 Code of Federal Regulations.

We (I) hereby claim foreign priority benefits under 35 U.S.C. § 119(a)–(d) or § 365(b) of any foreign application(s) for patent or inventor's certificate, or § 365(a) of any PCT International application which designated at least one country other than the United States, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate, or PCT International application having a filing date before that of the application on which priority is claimed. Prior Foreign Application(s)

Application No.	Country	Day/Month/Year	Priority Claimed
99106656.4	Europe	01 April 1999	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
99111321.8	Europe	10 June 1999	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
60/180687	United States of America	07 February 2000	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No

0093/000003

1-Anders Dahlqvist  
NAME OF SOLE OR FIRST INVENTOR

Anders Dahlqvist  
Signature of Inventor

Date 29.08.2001

Residence:  
Hemmansvägen 2  
24466 Furulund **SEX**  
Sweden  
Citizen of: Sweden  
Post Office Address: same as residence

2-Ulf Stahl  
NAME OF SECOND JOINT INVENTOR

Ulf Stahl  
Signature of Inventor

Date 29.08.2001

Residence:  
Liljegatan 7b  
75324 Uppsala **SEX**  
Sweden  
Citizen of: Sweden  
Post Office Address: same as residence

3-Marit Lenman  
NAME OF THIRD JOINT INVENTOR

Marit Lenman  
Signature of Inventor

Date 29.08.2001

Residence:  
Revingegatan 13a  
22359 Lund  
Sweden **SEX**  
Citizen of: Sweden  
Post Office Address: same as residence

4-Antoni Banas  
NAME OF FOURTH JOINT INVENTOR

Antoni Banas  
Signature of Inventor

Date 29.08.2001

Residence:  
Wiolinowa 14  
08110 Siedlce **PL**  
Poland  
Citizen of: ~~Sweden~~ **Poland**  
Post Office Address: same as residence

5-Hans Ronne  
NAME OF FIFTH JOINT INVENTOR

Hans Ronne  
Signature of Inventor

Date 29.08.2001


Residence:  
Dirigentvägen 169  
75654 Uppsala **SEX**  
Sweden  
Citizen of: Sweden  
Post Office Address: same as residence

We (I) hereby claim the benefit under Title 35, United States Codes, § 119(e) of any United States provisional application(s) listed below.

_____ (Application Number)	_____ (Filing Date)
_____ (Application Number)	_____ (Filing Date)

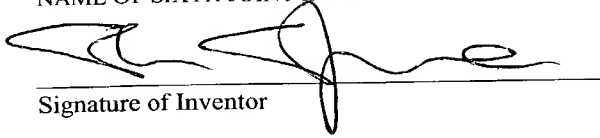
We (I) hereby claim the benefit under 35 U.S.C. § 120 of any United States application(s), or § 365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. § 112, I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR § 1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application.

Application Serial No.	Filing Date	Status (pending, patented, abandoned)
_____	_____	_____
_____	_____	_____
_____	_____	_____

 And we (I) hereby appoint **HERBERT. B. KEIL**, Registration Number 18,967; and **RUSSEL E. WEINKAUF**, Registration Number 18,495; the address of both being Messrs. Keil & Weinkauff, 1101 Connecticut Ave., N.W., Washington, D.C. 20036 (telephone 202-659-0100), our attorneys, with full power of substitution and revocation, to prosecute this application, to make alterations and amendments therein, to sign the drawings, to receive the patent, and to transact all business in the Patent Office connected therewith.

We (I) declare that all statements made herein of our (my) own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

6-10  
Sten Stymne  
NAME OF SIXTH JOINT INVENTOR

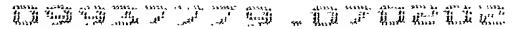
  
Signature of Inventor

Date 29.08.2001

Residence:  
Torrlösa 1380  
26990 Svalöv  
Sweden  
Citizen of: Sweden  
Post Office Address: same as residence

SEX





atg Met 1	ggc Gly	aca Thr	ctg Leu	ttt Phe 5	cga Arg	aga Arg	aat Asn	gtc Val	cag Gln 10	aac Asn	caa Gln	aag Lys	agt Ser	gat Asp 15	tct Ser	48
gat Asp	gaa Glu	aac Asn	aat Asn 20	aaa Lys	ggg Gly	ggt Gly	tct Ser	gtt Val 25	cat His	aac Asn	aag Lys	cga Arg	gag Glu 30	agc Ser	aga Arg	96
aac Asn	cac His	att Ile 35	cat His	cat His	caa Gln	cag Gln	gga Gly 40	tta Leu	ggc Gly	cat His	aag Lys	aga Arg 45	aga Arg	agg Arg	ggt Gly	144
att Ile	agt Ser 50	ggc Gly	agt Ser	gca Ala	aaa Lys	aga Arg 55	aat Asn	gag Glu	cgt Arg	ggc Gly	aaa Lys 60	gat Asp	ttc Phe	gac Asp	agg Arg	192
aaa Lys 65	aga Arg	gac Asp	ggg Gly	aac Asn	ggg Gly 70	aga Arg	aaa Lys	cgt Arg	tgg Trp	aga Arg 75	gat Asp	tcc Ser	aga Arg	aga Arg	ctg Leu 80	240
att Ile	ttc Phe	att Ile	ctt Leu	ggg Gly 85	gca Ala	ttc Phe	tta Leu	ggt Gly	gta Val 90	ctt Leu	ttg Leu	cct Pro	ttt Phe	agc Ser 95	ttt Phe	288
ggc Gly	gct Ala	tat Ile	cat Leu	gtt Gly	cat Leu	aat Asn	agc Gly	gat Val	agc Gly	gac Gly	ttg Leu	ttt Leu	gac Gly	aac Asn	ttt Gly	336

Gly	Ala	Tyr	His 100	Val	His	Asn	Ser	Asp 105	Ser	Asp	Leu	Phe	Asp 110	Asn	Phe	
gta Val	aat Asn	ttt Phe 115	gat Asp	tca Ser	ctt Leu	aaa Lys	gtg Val 120	tat Tyr	ttg Leu	gat Asp	gat Asp	tgg Trp 125	aaa Lys	gat Asp	gtt Val	384
ctc Leu	cca Pro 130	caa Gln	ggg Gly	ata Ile	agt Ser	tcg Ser 135	ttt Phe	att Ile	gat Asp	gat Asp	att Ile 140	cag Gln	gct Ala	ggg Gly	aac Asn	432
tac Tyr 145	tcc Ser	aca Thr	tct Ser	tct Ser	tta Leu 150	gat Asp	gat Asp	ctc Leu	agt Ser	gaa Glu 155	aat Asn	ttt Phe	gcc Ala	gtt Val	ggg Gly 160	480
aaa Lys	caa Gln	ctc Leu	tta Leu	cgt Arg 165	gat Asp	tat Tyr	aat Asn	atc Ile	gag Glu 170	gcc Ala	aaa Lys	cat His	cct Pro	gtt Val 175	gta Val	528
atg Met	gtt Val	cct Pro	ggg Gly 180	gtc Val	att Ile	tct Ser	acg Thr	gga Gly 185	att Ile	gaa Glu	agc Ser	tgg Trp	gga Gly 190	gtt Val	att Ile	576
gga Gly	gac Asp	gat Asp 195	gag Glu	tgc Cys	gat Asp	agt Ser	tct Ser 200	gcg Ala	cat His	ttt Phe	cgt Arg 205	aaa Lys 205	cgg Arg	ctg Leu	tgg Trp	624
gga Gly	agt Ser 210	ttt Phe	tac Tyr	atg Met	ctg Leu	aga Arg 215	aca Thr	atg Met	gtt Val	atg Met	gat Asp 220	aaa Lys	gtt Val	tgt Cys	tgg Trp	672
ttg Leu 225	aaa Lys	cat His	gta Val	atg Met	tta Leu 230	gat Asp	cct Pro	gaa Glu	aca Thr	ggg Gly 235	ctg Leu	gac Asp	cca Pro	ccg Pro	aac Asn 240	720
ttt Phe	acg Thr	cta Leu	cgt Arg	gca Ala 245	gca Ala	cag Gln	ggc Gly	ttc Phe	gaa Glu 250	tca Ser	act Thr	gat Asp	tat Tyr	ttc Phe 255	atc Ile	768
gca Ala	ggg Gly	tat Tyr 260	tgg Trp	att Ile	tgg Trp	aac Asn	aaa Lys	gtt Val 265	ttc Phe	caa Gln	aat Asn	ctg Leu 270	gga Gly 270	gta Val	att Ile	816
ggc Gly	tat Tyr	gaa Glu 275	ccc Pro	aat Asn	aaa Lys	atg Met	acg Thr 280	agt Ser	gct Ala	gcg Ala	tat Tyr	gat Asp 285	tgg Trp	agg Arg	ctt Leu	864
gca Ala	tat Tyr 290	tta Leu	gat Asp	cta Leu	gaa Glu	aga Arg 295	cgc Arg	gat Asp	agg Arg	tac Tyr	ttt Phe 300	acg Thr	aag Lys	cta Leu	aag Lys	912
gaa Glu 305	caa Gln	atc Ile	gaa Glu	ctg Leu	ttt Phe 310	cat His	caa Gln	ttg Leu	agt Ser	ggg Gly 315	gaa Glu	aaa Lys	gtt Val	tgt Cys	tta Leu 320	960
att Ile	gga Gly	cat His	tct Ser	atg Met 325	ggg Gly	tct Ser	cag Gln	att Ile	atc Ile 330	ttt Phe	tac Tyr	ttt Phe	atg Met 335	aaa Lys	tgg Trp	1008

gtc gag gct gaa ggc cct ctt tac ggt aat ggt ggt cgt ggc tgg gtt	1056
Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val	
340 345 350	
aac gaa cac ata gat tca ttc att aat gca gca ggg acg ctt ctg ggc	1104
Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly	
355 360 365	
gct cca aag gca gtt cca gct cta att agt ggt gaa atg aaa gat acc	1152
Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr	
370 375 380	
att caa tta aat acg tta gcc atg tat ggt ttg gaa aag ttc ttc tca	1200
Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser	
385 390 395 400	
aga att gag aga gta aaa atg tta caa acg tgg ggt ggt ata cca tca	1248
Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser	
405 410 415	
atg cta cca aag gga gaa gag gtc att tgg ggg gat atg aag tca tct	1296
Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser	
420 425 430	
tca gag gat gca ttg aat aac aac act gac aca tac ggc aat ttc att	1344
Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile	
435 440 445	
cga ttt gaa agg aat acg agc gat gct ttc aac aaa aat ttg aca atg	1392
Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met	
450 455 460	
aaa gac gcc att aac atg aca tta tcg ata tca cct gaa tgg ctc caa	1440
Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln	
465 470 475 480	
aga aga gta cat gag cag tac tcg ttc ggc tat tcc aag aat gaa gaa	1488
Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu	
485 490 495	
gag tta aga aaa aat gag cta cac cac aag cac tgg tcg aat cca atg	1536
Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met	
500 505 510	
gaa gta cca ctt cca gaa gct ccc cac atg aaa atc tat tgt ata tac	1584
Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr	
515 520 525	
ggg gtg aac aac cca act gaa agg gca tat gta tat aag gaa gag gat	1632
Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp	
530 535 540	
gac tcc tct gct ctg aat ttg acc atc gac tac gaa agc aag caa cct	1680
Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro	
545 550 555 560	
gta ttc ctc acc gag ggg gac gga acc gtt ccg ctc gtg gcg cat tca	1728
Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser	
565 570 575	

atg tgt cac aaa tgg gcc cag ggt gct tca ccg tac aac cct gcc gga 1776  
Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly  
580 585 590

att aac gtt act att gtg gaa atg aaa cac cag cca gat cga ttt gat 1824  
Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp  
595 600 605

ata cgt ggt gga gca aaa agc gcc gaa cac gta gac atc ctc ggc agc 1872  
Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser  
610 615 620

gcg gag ttg aac gat tac atc ttg aaa att gca agc ggt aat ggc gat 1920  
Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp  
625 630 635 640

ctc gtc gag cca cgc caa ttg tct aat ttg agc cag tgg gtt tct cag 1968  
Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln  
645 650 655

atg ccc ttc cca atg taa 1986  
Met Pro Phe Pro Met  
660

<210> 2

<211> 661

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 2

Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser  
1 5 10 15

Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg  
20 25 30

Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly  
35 40 45

Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg  
50 55 60

Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu  
65 70 75 80

Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe  
85 90 95

Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe  
100 105 110

Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val  
115 120 125

Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn  
130 135 140



Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln  
 465 470 475 480

Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu  
 485 490 495

Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met  
 500 505 510

Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr  
 515 520 525

Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp  
 530 535 540

Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro  
 545 550 555 560

Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser  
 565 570 575

Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly  
 580 585 590

Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp  
 595 600 605

Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser  
 610 615 620

Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp  
 625 630 635 640

Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln  
 645 650 655

Met Pro Phe Pro Met  
 660

<210> 3  
 <211> 2312  
 <212> DNA  
 <213> Schizosaccharomyces pombe  
 <400> 3

atggcgtctt ccaagaagag caaaactcat aagaaaaaga aagaagtcaa atctcctatc 60  
 gacttaccaa attcaaagaa accaactcgc gctttgagtg agcaaccttc agcgtccgaa 120  
 acacaatctg tttcaaataa atcaagaaaa tctaaatttg gaaaaagatt gaattttata 180  
 ttggggcgcta ttttggaat atgcggtgct ttttttttcg ctgttgagga cgacaatgct 240  
 gttttcgacc ctgctacggt agataaattt ggggaatatgc taggctcttc agacttgttt 300  
 gatgacatta aaggatattt atcttataat gtgtttaagg atgcaccttt tactacggac 360

aagccttcgc agtctcctag cggaaatgaa gttcaagttg gtcttgatat gtacaatgag 420  
ggatategaa gtgaccatcc tgttattatg gttcctggtg ttatcagctc aggattagaa 480  
agttggtcgt ttaataattg ctcgattcct tacttttagga aacgtctttg gggtagctgg 540  
tctatgctga aggcaatggt ccttgacaag caatgctggc ttgaacattt aatgcttgat 600  
aaaaaaaccg gcttgatcc gaagggaatt aagctgcgag cagctcaggg gtttgaagca 660  
gctgattttt ttatcacggg ctattggatt tggagtaaag taattgaaaa ccttgctgca 720  
attggttatg agcctaataa catgttaagt gcttcctacg attggcgggt atcatatgca 780  
aatttagagg aacgtgataa atatttttca aagttaaaaa tgttcattga gtacagcaac 840  
attgtacata agaaaaagg agtggtgatt tctcactcca tgggttcaca gggtacgtac 900  
tattttttta agtgggttga agctgagggc tacggaaatg gtggaccgac ttgggttaat 960  
gatcatattg aagcatttat aaatgtgagt ctgatgggt gtttgactac gtttctaact 1020  
tttgaataga tatcgggatc tttgattgga gcacccaaaa cagtggcagc gcttttatcg 1080  
ggtgaaatga aagatacagg tattgtaatt acattaaaca tgttaatat taatttttgc 1140  
taaccgtttt aagctcaatt gaatcagttt tcggtctatg ggtaagcaat aaattgttga 1200  
gatttggttac taatttactg tttagtttgg aaaaattttt ttcccgttct gaggtatat 1260  
caaaaataca aatgtgctct actttttcta acttttaata gagagccatg atgggttcgca 1320  
ctatgggagg agttagtctt atgcttccta aaggaggcga tgttgatgg ggaaatgcc 1380  
gttgggtaag aaatatgtgc tgttaatttt ttattaatat ttaggctcca gatgatctta 1440  
atcaaacaaa tttttccaat ggtgcaatta ttgatatag agaagacatt gataaggacc 1500  
acgatgaatt tgacatagat gatgcattac aatttttaaa aaatgttaca gatgacgatt 1560  
ttaaagtcac gctagcgaaa aattattccc acggtcttgc ttggactgaa aaagaagtgt 1620  
taaaaaataa cgaaatgccg tctaaatgga taaatccgct agaagtaaga acattaaagt 1680  
tactaaatta tactaaccca aatagactag tcttccttat gtcctgata tgaaaattta 1740  
ttgcgttcac ggggtcggaa aaccaactga gagagggtat tattatacta ataatcctga 1800  
ggggcaacct gtcattgatt cctcggttaa tgatggaaca aaagttgaaa atgtgagaga 1860  
atztatgttt caaacattct attactggt ttattaggg attgttatgg atgatggtga 1920  
tggaacttta ccaatattag cccttggttt ggtgtgcaat aaagtttggc aaacaaaaag 1980  
gtttaatcct gctaatacaa gtatcacaaa ttatgaaatc aagcatgaac ctgctgcgtt 2040  
tgatctgaga ggaggacctc gctcggcaga acacgtcgat atacttgac attcagagct 2100  
aaatgtatgt tcattttacc ttacaaattt ctattactaa ctcttgaaat aaggaaatta 2160

ttttaaaggt ttcattcaggc catgggtgact cggtagcaaaa ccgttatata tcagatatcc 2220  
 agtacggaca taagttttgt agattgcaat taactaacta accgaacagg gaaataataa 2280  
 atgagataaa tctcgataaa cctagaaatt aa 2312

<210> 4  
 <211> 3685  
 <212> DNA  
 <213> Arabidopsis thaliana

<400> 4

atgccccctta ttcattcggaa aaagccgacg gagaaaccat cgacgcccgc atctgaagag 60  
 gtggtgcacg atgaggattc gcaaaagaaa ccacacgaat cttccaaatc ccaccataag 120  
 aaatcgaacg gaggaggga ggtggtcgtgc atcgattctt gttggttggt cattgggtgt 180  
 gtgtgtgtaa cctggtggtt tcttctcttc ctttacaacg caatgcctgc gagcttcctt 240  
 cagtatgtaa cggagcgaat cacgggtcct ttgcctgacc cgcccgggtg taagctcaaa 300  
 aaagaaggtc ttaaggcgaa acatcctggt gtcttcattc ctgggattgt caccgggtggg 360  
 ctcgagcttt gggaaggcaa acaatgcgct gatggtttat ttagaaaacg tttgtggggg 420  
 ggaacttttg gtgaagtcta caaaagggtg gctcaacaat tctcactctt cttttatatt 480  
 gggatttgga ttggatctga tgagatcacg cacttggtgc ttcttcaaca tcactcaaac 540  
 ttttaattcca tgtttgtctg tcttactctt tacttttttt tttttttgat gtgaaacgct 600  
 attttcttaa gagactatct ctgtatgtgt aaggtaagcg ttccaaggac gtaattgggt 660  
 tggactatct ctgtttgatt gttaacttta ggatataaaa tagctgcctt ggaatttcaa 720  
 gtcattctat tgccaaatct gttgctagac atgcccctaga gtccgttcat aacaagttac 780  
 ttcctttact gtcgttgctg gtagatttag ctttgtgtag cgtataatga agtagtggtt 840  
 tatgttttgt tgggaataga gaagttctaa ctacatctgt ggaaagtgtg ttcaggctgt 900  
 gatagaggac tgttgcttta ttattcaact atgtatatgt gtaattaaag ctagttcctt 960  
 tttgatcttt cagctcaatg tgcttttctc aatttttttc tcaatttcaa agtttcacat 1020  
 cgagtttatt cacatgtctt gaatttgcgc catcctcggt ctgttatcca gctttgaact 1080  
 cctcccgacc ctgctatgga tatattaaaa aaaaagtgtt ttgtgggttg catctttgtt 1140  
 acgatctgca tcttcttctt tcggctcagt gttcatgttt ttgctatggt agagatgggc 1200  
 aatgttattg ttgatggtaa cagtgggtata gttgatagta tcttaactaa tcaattatct 1260  
 ctttgattca ggcctctatg ttgggtggaa cacatgtcac ttgacaatga aactgggttg 1320  
 gatccagctg gtattagagt tcgagctgta tcaggactcg tggctgctga ctactttgct 1380



cctggctact ttgtctgggc agtgctgatt gctaaccttg cacatattgg atatgaagag 1440  
 aaaaatatgt acatggctgc atatgactgg cggttttctg ttcagaacac agaggttctt 1500  
 ttctcatcgt tctttctatt attctgttcc atgttacggt tctttcttca ttacttaagg 1560  
 cttaaatatg tttcatgttg aattaatagg tacgtgatca gactcttagc cgtatgaaaa 1620  
 gtaatataga gttgatgggt tctaccaacg gtggaaaaaa agcagttata gttccgcatt 1680  
 ccatgggggt cttgtatttt ctacatttta tgaagtgggt tgaggcacca gtcctctctg 1740  
 gtggcggggg tgggccagat tgggtgtgcaa agtatattaa ggcggtgatg aacattgggtg 1800  
 gaccatttct tgggtgttcca aaagctgttg cagggtcttt ctctgctgaa gcaaaggatg 1860  
 ttgcagttgc cagggtattga atatctgctt atacttttga tgatcagaac cttggctctg 1920  
 gaactcaaag ttattctact aaatatcaat tctaataaca ttgctatatt atcgtgcaa 1980  
 ctgacattgg ttgattatth ttgctgctta tgtaactgaa actctcttga gattagacaa 2040  
 atgatgaatt gataattctt acgcattgct ctgtgatgac cagtttctta gcttcgacga 2100  
 taacatttgt catactgtct tttggagggc attgaatttt gctatggaaa gcgctggagc 2160  
 ttccatgctt gcattcttta ccaattagcg ttattctgct tctttcaatt ttcttgtata 2220  
 tgcactctat gtcttttatt tctttttaat taaagactcg ttggattagt tgctctatta 2280  
 gtcacttggg tctttaatat agaactttac tttcttcgaa aattgcagag cgattgcccc 2340  
 aggattctta gacaccgata tatttagact tcagaccttg cagcatgtaa tgagaatgac 2400  
 acgcacatgg gactcaacaa tgtctatggt accgaaggga ggtgacacga tatggggcgg 2460  
 gcttgattgg tcaccggaga aaggccacac ctgttgtggg aaaaagcaaa agaacaacga 2520  
 aacttgtggg gaagcagggtg aaaacggagt ttccaagaaa agtcctgtta actatggaag 2580  
 gatgatattt tttgggaaag aagtagcaga ggctgcgcca tctgagatta ataattattga 2640  
 ttttcgagta aggacatata aatcataata aaccttgtag attttgtgat tgtatgatga 2700  
 atatctgtac attttatctg gtgaagggtg ctgtcaaagg tcagagtatc ccaaatcaca 2760  
 cctgtcgtga cgtgtggaca gagtaccatg acatgggaat tgctgggatc aaagctatcg 2820  
 ctgagtataa ggtctacact gctggtgaag ctatagatct actacattat gttgctccta 2880  
 agatgatggc gcgtggtgcc gctcatttct cttatggaat tgctgatgat ttggatgaca 2940  
 ccaagtatca agatcccaaa tactgggtcaa atccgttaga gacaaagtaa gtgatttctt 3000  
 gattccaact gtatccttcg tctgatgca ttatcagctt tttgttttc ggtcttggtg 3060  
 gatatggttt tcagctcaaa gcttacaaag ctgtttctga gcctttctca aaaaggcttg 3120  
 ctgagtaata ttgaggtgct aaagttgata catgtgactc ttgcttataa atcctccggt 3180

tggtttgttc tgctttttca gattaccgaa tgctcctgag atggaaatct actcattata 3240  
 cggagtgggg ataccaacgg aacgagcata cgtatacaag cttaaccagt ctcccgacag 3300  
 ttgcatcccc ttccagatat tcactttctgc tcacgaggag gacgaagata gctgtctgaa 3360  
 agcaggagtt tacaatgtgg atggggatga aacagtaccc gtcctaagtg ccgggtacat 3420  
 gtgtgcaaaa gcgtggcgtg gcaagacaag attcaaccct tccggaatca agacttatat 3480  
 aagagaatac aatcactctc cgccggctaa cctgttgga gggcgcgga cgagagtg 3540  
 tgcccatgtt gatatcatgg gaaactttgc tttgatcgaa gatatcatga gggttgccgc 3600  
 cggaggtaac gggctctgata taggacatga ccagggtcac tctggcatat ttgaatggtc 3660  
 ggagcgtatt gacctgaagc tgtga 3685

<210> 5  
 <211> 2427  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> unsure  
 <222> 1..2427  
 <223> n = a or g or c or t/u

<400> 5

agaaacagct ctttgtctct ctcgactgat ctaacaatcc ctaatctgtg ttctaaattc 60  
 ctggacgaga ttgacaaaag tccgtatagc ttaacctggg ttaatttcaa gtgacagata 120  
 tgcccccttat tcatcggaag aagccgacgg agaaaccatc gacgccgcca tctgaagagg 180  
 tgggtgcacga tgaggattcg caaaagaaac cacacgaatc ttccaaatcc caccataaga 240  
 aatcgaacgg aggagggaag tggtcgtgca tcgattcttg ttgttggttc attgggtgtg 300  
 tgtgtgtaac ctggtggttt cttctcttcc tttacaacgc aatgcctgag agcttccctc 360  
 agtatgtaac ggagcgaatc acgggtcctt tgccctgacc gcccggtgtt aagctcaaaa 420  
 aaagaaggtc ttaaggcgaa acatcctgtt gtcttcattc ctgggattgt caccgggtggg 480  
 ctcgagcttt gggaaggcaa acaatgcgct gatgggttat ttagaaaacg tttgtggggg 540  
 ggaacttttg gtgaagtcta caaaaggcct ctatgttggg tggaacacat gtcacttgac 600  
 aatgaaactg ggttgatcc agctgggtatt agagtccgag ctgtatcagg actcgtggct 660  
 gctgactact ttgctcctgg ctactttgtc tgggcagtgc tgattgctaa ccttgacat 720  
 attggatatg aagagaaaaa tatgtacatg gctgcatatg actggcggct ttcgtttcag 780  
 aacacagagg tacgtgatca gactottagc cgtatgaaaa gtaatataga gttgatgggt 840  
 tctaccaacg gtggaaaaaa agcagttata gttccgcatt ccatgggggt cttgtatttt 900

11

```

ctacatttta tgaagtgggt tgaggcacca gtcctctctg gtggcggggg tgggccagat 960
tgggtgtgcaa agtatattaa ggcgggtgatg aacattgggtg gaccatttct tgggtgttcca 1020
aaagctgttg cagggccttt ctctgctgaa gcaaaggatg ttgcagttgc cagagcgatt 1080
gccccaggat tcttagacac cgatatatct agacttcaga ccttgacgca tgtaatgaga 1140
atgacacgca catgggactc aacaatgtct atgttaccga agggaggtga cacgatatgg 1200
ggcgggcttg attggtcacc ggagaaaggc cacacctgtt gtgggaaaaa gcaaaagaac 1260
aacgaaactt gtgggtgaagc aggtgaaaac ggagtttcca agaaaagtcc tgttaactat 1320
ggaaggatga tatcttttgg gaaagaagta gcagaggctg cgccatctga gattaataat 1380
attgattttc gaggtgctgt caaaggctcag agtatcccaa atcacacctg tcgtgacgtg 1440
tggacagagt accatgacat ggaattgtct gggatcaaag ctatcgctga gtataaggtc 1500
tacctgctg gtgaagctat agatctacta cattatgttg ctctaagat gatggcgct 1560
ggtgcgctc atttctctta tggaattgtc gatgatttgg atgacaccaa gtatcaagat 1620
cccaaatact ggtcaaactc gtagagaca aaattaccga atgctcctga gatggaaatc 1680
tactcattat acggagtggg gataccaacg gaacgagcat acgtatacaa gcttaaccag 1740
tctcccgaca gttgcatccc ctttcagata ttcacttctg ctcacgagga ggacgaagat 1800
agctgtctga aagcaggagt ttacaatgtg gatggggatg aaacagtacc cgtcctaagt 1860
gccgggtaca tgtgtgcaaa agcgtggcgt ggcaagacaa gattcaacct ttccggaatc 1920
aagacttata taagagaata caatcactct ccgccggcta acctgttgga agggcgcggg 1980
acgcagagtg gtgcccattg tgatatcatg ggaaactttg ctttgatcga agatatcatg 2040
aggggttgccg ccggaggtaa cgggtctgat ataggacatg accaggtcca ctctggcata 2100
tttgaatggc cggagcgtat tgacctgaag ctgtgaatat catgatctct ttaagctgtc 2160
ctgtcagctt atgtgaatcc aatactttga aagagagatc atcatcaatt catcatcatc 2220
gtcatcatca tgatgtctca ctcacaaaga agcctgagaa tgatactttg gtgcgaaatt 2280
ctcaatacct cttaaatatt cttattgaat gtaaattata caatcctatc taatgtttga 2340
acgataacac aaaacttgct gcngccatgt ttgtttgtct tgtcaaaagc atcaatttgt 2400
gggttaaaaa aaaaaaaaaa aaaaaaa 2427

```

```

<210> 6
<211> 671
<212> PRT
<213> Arabidopsis thaliana
<400> 6

```



Ala Pro Gly Phe Leu Asp Thr Asp Ile Phe Arg Leu Gln Thr Leu Gln  
 325 330 335  
 His Val Met Arg Met Thr Arg Thr Trp Asp Ser Thr Met Ser Met Leu  
 340 345 350  
 Pro Lys Gly Gly Asp Thr Ile Trp Gly Gly Leu Asp Trp Ser Pro Glu  
 355 360 365  
 Lys Gly His Thr Cys Cys Gly Lys Lys Gln Lys Asn Asn Glu Thr Cys  
 370 375 380  
 Gly Glu Ala Gly Glu Asn Gly Val Ser Lys Lys Ser Pro Val Asn Tyr  
 385 390 395 400  
 Gly Arg Met Ile Ser Phe Gly Lys Glu Val Ala Glu Ala Ala Pro Ser  
 405 410 415  
 Glu Ile Asn Asn Ile Asp Phe Arg Gly Ala Val Lys Gly Gln Ser Ile  
 420 425 430  
 Pro Asn His Thr Cys Arg Asp Val Trp Thr Glu Tyr His Asp Met Gly  
 435 440 445  
 Ile Ala Gly Ile Lys Ala Ile Ala Glu Tyr Lys Val Tyr Thr Ala Gly  
 450 455 460  
 Glu Ala Ile Asp Leu Leu His Tyr Val Ala Pro Lys Met Met Ala Arg  
 465 470 475 480  
 Gly Ala Ala His Phe Ser Tyr Gly Ile Ala Asp Asp Leu Asp Asp Thr  
 485 490 495  
 Lys Tyr Gln Asp Pro Lys Tyr Trp Ser Asn Pro Leu Glu Thr Lys Leu  
 500 505 510  
 Pro Asn Ala Pro Glu Met Glu Ile Tyr Ser Leu Tyr Gly Val Gly Ile  
 515 520 525  
 Pro Thr Glu Arg Ala Tyr Val Tyr Lys Leu Asn Gln Ser Pro Asp Ser  
 530 535 540  
 Cys Ile Pro Phe Gln Ile Phe Thr Ser Ala His Glu Glu Asp Glu Asp  
 545 550 555 560  
 Ser Cys Leu Lys Ala Gly Val Tyr Asn Val Asp Gly Asp Glu Thr Val  
 565 570 575  
 Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Ala Trp Arg Gly Lys  
 580 585 590  
 Thr Arg Phe Asn Pro Ser Gly Ile Lys Thr Tyr Ile Arg Glu Tyr Asn  
 595 600 605  
 His Ser Pro Pro Ala Asn Leu Leu Glu Gly Arg Gly Thr Gln Ser Gly  
 610 615 620  
 Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile Glu Asp Ile Met

His Ser Gly Ile Phe Glu Trp Ser Glu Arg Ile Asp Leu Lys Leu  
660 665 670

```
<220>
<221> CDS
<222> (1)..(402)
```

```
<221> unsure
<222> 1..643
<223> n= a or q or c or t/u
```

cgg gag aaa ata gct gct ttg aag ggg ggt gtt tac tta gcc gat ggt 48  
Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly  
1 5 10 15

gat gaa act gtt cca gtt ctt agt gcg ggc tac atg tgt gcg aaa gga 96  
Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly  
20 25 30

tgg	cgt	ggc	aaa	act	cgt	ttc	agc	cct	gcc	ggc	agc	aag	act	tac	gtg	144
Trp	Arg	Gly	Lys	Thr	Arg	Phe	Ser	Pro	Ala	Gly	Ser	Lys	Thr	Tyr	Val	
		35					40					45				

aga gaa tac agc cat tcg cca ccc tct act ctc ctg gaa ggc agg ggc 192  
Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly  
50 55 60

acc cag agc ggt gca cat gtt gat ata atg ggg aac ttt gct cta att 240  
Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile  
65 70 75 80

gag gac gtc atc aga ata gct gct ggg gca acc ggt gag gaa att ggt 288  
Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly  
85 90 95

ggc gat cag gtt tat tca gat ata ttc aag tgg tca gag aaa atc aaa 336  
Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Lys  
100 105 110

ttg aaa ttg taa cct atg gga agt taa aga agt gcc gac ccg ttt att 384  
Leu Lys Leu  
115

gcg ttc caa agt gtc ctg cctgagtgca actctggatt ttgcttaaatt 432

attgtaattt ttcacgcttc attcgtccct ttgtcaaatt tacatttgac aggacgccaa 492

15

tgcgatacga tgttgtagcg ctattttcag cattgtatat taaactgtac aggtgtaagt 552  
tgcatttgcc agctgaaatt gtgtagtcgt tttctttacg atttaaatanc aagtggcgga 612  
gcagtgcccc aagcnaaaaa aaaaaaaaaa a 643

<210> 8  
<211> 115  
<212> PRT  
<213> Zea mays

<400> 8

Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly  
1 5 10 15  
Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly  
20 25 30  
Trp Arg Gly Lys Thr Arg Phe Ser Pro Ala Gly Ser Lys Thr Tyr Val  
35 40 45  
Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly  
50 55 60  
Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile  
65 70 75 80  
Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly  
85 90 95  
Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Lys  
100 105 110  
Leu Lys Leu  
115

<210> 9  
<211> 616  
<212> DNA  
<213> Neurospora crassa

<220>  
<221> unsure  
<222> 1..616  
<223> n= a or g or c or t/u

<400> 9

ggtggcgaag acganggcgg aagttggagg ctaacgagaa tgacnctcgg agatggatct 60  
accctctaga gacacgacta ccnttgccacc cagcctcaag gtntacngtt tntatgggta 120  
ggaagccgac ggagcgagcc tacatctatc tggcgccccga tcccgggacg acaacgcac 180  
tttagatgac gatcgatacg actttgactn aggggcacat tgaccacggg gtgattttgg 240  
gcgaaggcga tggcacagtg aaccttatga gtttggggta cctgtgcaat aaggggtgga 300

aaatgaagag atacaatcct gcgggctcaa aaataaccgt ggtcgagatg ccgcatgaac 360  
 cagaacgggt caatccgaga ggaggccga atacggcgga tcacgtggat attctaggaa 420  
 ggcagaatct aaacgagtac attcttaaag tggcggcagg tcgaggcgat acaattgagg 480  
 attttattac tagtaatatt cttaaataatg tagaaaagggt tgaaatttat gaagagtaat 540  
 taaatacggc acatagggtta ctcaatagta tgactaatta aaaaaaaatt ttttttctaa 600  
 aaaaaaaaaa aaaaaa 616

<210> 10  
 <211> 1562  
 <212> DNA  
 <213> *Arabidopsis thaliana*

<400> 10

atgaaaaaaaa tatcttcaca ttattcggta gtcatacgga tactcgttgt ggtgacgatg 60  
 acctcgatgt gtcaagctgt gggtagcaac gtgtaccctt tgattctgggt tccaggaaac 120  
 ggaggtaacc agctagaggt acggctggac agagaatata agccaagtag tgtctgggtgt 180  
 agcagctgggt tatatccgat tcataagaag agtggtggat ggtttaggct atggttcgat 240  
 gcagcagtgt tattgtctcc cttcaccagg tgcttcagcg atcgaatgat gttgtactat 300  
 gaccctgatt tggatgatta ccaaaatgct cctgggtgtcc aaaccgggt tectcatttc 360  
 ggttcgacca aatcacttct atacctcgac cctcgtctcc ggtagtact ttccaagata 420  
 tatcattttg ggacatttgc ataatagaaca aaatagacat aaatttgggg gattattgtt 480  
 atatcaatat ccatttatat gctagtcgggt aatgtgagtg ttatgttagt atagttaatg 540  
 tgagtgttat gtgattttcc attttaaatg aagctagaaa gttgtcgttt aataatgttg 600  
 ctatgtcatg agaattataa ggacactatg taaatgtagc ttaataataa ggtttgattt 660  
 gcagagatgc cacatcttac atggaacatt tggtgaaagc tctagagaaa aaatgcgggt 720  
 atgttaacga ccaaaccatc ctaggagctc catatgattt caggtagcgc ctggctgctt 780  
 cgggccaccc gtcccgtgta gcctcacagt tectacaaga cctcaaaca ttggtggaaa 840  
 aaactagcag cgagaacgaa ggaaagccag tgatactcct ctcccatagc ctaggaggac 900  
 ttttcgtcct ccatttcctc aaccgtacca ccccttcatt gcgcgcgaag tacatcaaac 960  
 actttgttgc actcgtcgcg ccatgggggtg ggacgatctc tcagatgaag acatttgctt 1020  
 ctggcaacac actcgtgtgc cctttagtta accctttgct ggtcagacgg catcagagga 1080  
 cctccgagag taaccaatgg ctacttccat ctaccaaagt gtttcacgac agaactaaac 1140  
 cgcttgctgt aactccccag gttaactaca cagcttacga gatggatcgg ttttttgcag 1200



atgggagcga	attcgaaatc	agtaacggct	tccttcaccg	tcatcgccgt	ttttttcttg	60
at ttg cgg tg	gc cga act gc	gg tgg agg at	gag acc gag t	tt cac ggc ga	ct act cga ag	120
ct atc ggg ta	ta atc att cc	gg gatt tgc g	tc gac gc agc	ta cgag cgt g	gt cga tcct t	180
gact gtcc at	ac act ccgt t	gg act tca at	cc gct cga cc	tc gtat ggct	ag acacc act	240
aagg tccg tg	at ctt catt t	cct tcc gct cc	tt att ctg tc	gg tcc gag tca	ct tgt ttgat g	300
aat tcca agc	gaa atata gc	aat gaag cat	gt ctccgt ctc	tct tatt gat	tc gtt catta	360
gt caac agt g	ac gct tct ga	at ctg agt tt	ag agt cat at	aaa acag ctg	act cgg cgag	420
tg ttt ccc at	cg ctt ttg gt	tc gc taa atg	tag cgc aat g	aat gtg taat	tag tct gcgc	480
tt ttt att ca	act agat ctg	ca agt ttt tc	ag agt gct ca	at agt agt ta	gaaa atgt ta	540
gg tcatt tta	ct tgt gcatt	gt gatt ctt t	tg gtt gtt gc	tt act gat cg	ac gtgat gga	600
tg gtt tac ag	ct tct ttt ctg	ct gtca act g	ct ggt tta ag	tgt atggt gc	tag atc ctt a	660
ta atcaa aca	gacc atccc g	agt gta agt c	ac ggc ctg ac	agt ggt ctt t	cag ccac tca c	720
aga attgg at	cc aggt taca	ta acagg tag	tt tccg att t	tt ctt tct t t	tg agt ttt ct	780
tca atttg at	at cat ctt gt	tgt gat ataa	tat ggc taag	tt cat taatt	tggt caatt t	840
tc aggtc ctc	tt tct act gt	ct gga aag ag	tg gct taag t	gg tgt gtt ga	gt ttg gtata	900
ga agcaa atg	ca att gtc gc	tgt tcc atac	gatt ggag at	tgt cac caac	caa attgg aa	960
gag cgtg acc	tt tact ttt ca	ca agt cta ag	tt agt cct ta	tc agg cta at	gt ctt ttt atc	1020
tt ctct ttt t	at gta agata	ag cta agag c	tct ggt cgt c	tt cct ttt tg	cagg ttg acc	1080
ttt gaa actg	ctt taaa act	ccgt ggc ggc	cct tct atag	tatt tgc cca	tt caat qgg t	1140



aggtaatttt ccgcaatggc agaagtaaaa caggaaggca aagtcttctg tatcagtcta 3000  
gtggcatgtt atctcagttg cataagcaaa ttattaaaca actaaaattt aagtactttt 3060  
ttatcattcc ttttgagctt agtggatgat cagtggctta aagtgggaag aggtgttgca 3120  
tgaaacatga cacttgtatc aaagataact agcaaaaaca aactaaccga tttctgaatt 3180  
tcatattatt aggagtagtc gtgcttttaa aaaatttggt ttaagaaacc gaaaaactag 3240  
ttcatatctt gattgtgcaa tatctgcagg tctggaactg tggttgatgg gaacgctgga 3300  
cctataactg gggatgagac ggtaagctca gaagttgggt ttgaaattat cttcttgcaa 3360  
actactgaag actaagataa tacttgcttc tggaacactg cttgctatgt tctctagtac 3420  
actgcaatat tgactctccg ctacttttat tgattatgaa attgatctct tataggtacc 3480  
ctatcattca ctctcttggg gcaagaattg gctcggacct aaagttaaca taacaatggc 3540  
tccccaggta ctctttttta gttcctcacc ttatatagat caaactttta gtgtactttt 3600  
ctgggttatgt gttgatttac ctccaatttg ttctttctaa aaatcatata tctctgtact 3660  
cctcaagaac ttgtattaat ctaaacgaga ttctcattgg gaaaataaaa caacagccag 3720  
aacacgatgg aagcgacgta catgtggaac taaatgttga tcatgagcat gggtcagaca 3780  
tcatagctaa catgacaaaa gcaccaaggg ttaagtacat aaccttttat gaagactctg 3840  
agagcattcc ggggaagaga accgcagctc gggagcttga taaaagtggg tattaa 3896

<210> 12  
<211> 709  
<212> DNA  
<213> *Lycopersicon esculentum*  
<400> 12

ctggggccaa aagtgaacat aacaaggaca ccacagtcag agcatgatgt tcagatgtac 60  
aagtgcattc aaatatagag catcaacatg gtgaagatat cattcccaat atgacaaaagt 120  
tacctacaat gaagtacata acctattatg aggattctga aagttttcca gggacaagaa 180  
cagcagtttg ggagcttgat aaagcaaata acaggaacat tgcagatct ccagctttga 240  
tgccgggagct gtggcttgag atgtggcatg atattcatcc tgataaaaag tccaagtttg 300  
ttacaaaagg tgggtgtctga tctcactat tttcttctat aaatgtttga gtttgtattg 360  
acattgtaag tattgcaaca aaaagcaaag cgtgggcctc tgagggatga ggactgctat 420  
tgggattacg ggaaagctcg atgtgcatgg gctgaacatt gtgaatacag gttagaatat 480  
tcaaattata ttttgcaaaa tattctcttt ttgtgtattt aggccacct tccccggtca 540



21

<210> 13  
 <211> 623  
 <212> PRT  
 <213> Schizosaccharomyces pombe

<400> 13

Met	Ala	Ser	Ser	Lys	Lys	Ser	Lys	Thr	His	Lys	Lys	Lys	Lys	Glu	Val
1				5					10					15	
Lys	Ser	Pro	Ile	Asp	Leu	Pro	Asn	Ser	Lys	Lys	Pro	Thr	Arg	Ala	Leu
			20					25					30		
Ser	Glu	Gln	Pro	Ser	Ala	Ser	Glu	Thr	Gln	Ser	Val	Ser	Asn	Lys	Ser
		35					40					45			
Arg	Lys	Ser	Lys	Phe	Gly	Lys	Arg	Leu	Asn	Phe	Ile	Leu	Gly	Ala	Ile
	50					55					60				
Leu	Gly	Ile	Cys	Gly	Ala	Phe	Phe	Phe	Ala	Val	Gly	Asp	Asp	Asn	Ala
65					70					75					80
Val	Phe	Asp	Pro	Ala	Thr	Leu	Asp	Lys	Phe	Gly	Asn	Met	Leu	Gly	Ser
				85					90					95	
Ser	Asp	Leu	Phe	Asp	Asp	Ile	Lys	Gly	Tyr	Leu	Ser	Tyr	Asn	Val	Phe
		100						105					110		
Lys	Asp	Ala	Pro	Phe	Thr	Thr	Asp	Lys	Pro	Ser	Gln	Ser	Pro	Ser	Gly
		115					120					125			
Asn	Glu	Val	Gln	Val	Gly	Leu	Asp	Met	Tyr	Asn	Glu	Gly	Tyr	Arg	Ser
	130					135					140				
Asp	His	Pro	Val	Ile	Met	Val	Pro	Gly	Val	Ile	Ser	Ser	Gly	Leu	Glu
145					150					155					160
Ser	Trp	Ser	Phe	Asn	Asn	Cys	Ser	Ile	Pro	Tyr	Phe	Arg	Lys	Arg	Leu
				165					170					175	
Trp	Gly	Ser	Trp	Ser	Met	Leu	Lys	Ala	Met	Phe	Leu	Asp	Lys	Gln	Cys
		180						185					190		
Trp	Leu	Glu	His	Leu	Met	Leu	Asp	Lys	Lys	Thr	Gly	Leu	Asp	Pro	Lys
		195					200					205			
Gly	Ile	Lys	Leu	Arg	Ala	Ala	Gln	Gly	Phe	Glu	Ala	Ala	Asp	Phe	Phe
	210					215					220				
Ile	Thr	Gly	Tyr	Trp	Ile	Trp	Ser	Lys	Val	Ile	Glu	Asn	Leu	Ala	Ala
225					230					235					240
Ile	Gly	Tyr	Glu	Pro	Asn	Asn	Met	Leu	Ser	Ala	Ser	Tyr	Asp	Trp	Arg
				245					250					255	
Leu	Ser	Tyr	Ala	Asn	Leu	Glu	Glu	Arg	Asp	Lys	Tyr	Phe	Ser	Lys	Leu
			260					265					270		
Lys	Met	Phe	Ile	Glu	Tyr	Ser	Asn	Ile	Val	His	Lys	Lys	Lys	Val	Val

275						280						285				
Leu	Ile	Ser	His	Ser	Met	Gly	Ser	Gln	Val	Thr	Tyr	Tyr	Phe	Phe	Lys	
	290					295					300					
Trp	Val	Glu	Ala	Glu	Gly	Tyr	Gly	Asn	Gly	Gly	Pro	Thr	Trp	Val	Asn	
305					310					315					320	
Asp	His	Ile	Glu	Ala	Phe	Ile	Asn	Ile	Ser	Gly	Ser	Leu	Ile	Gly	Ala	
				325					330					335		
Pro	Lys	Thr	Val	Ala	Ala	Leu	Leu	Ser	Gly	Glu	Met	Lys	Asp	Thr	Gly	
			340					345					350			
Ile	Val	Ile	Thr	Leu	Asn	Ile	Leu	Glu	Lys	Phe	Phe	Ser	Arg	Ser	Glu	
		355					360					365				
Arg	Ala	Met	Met	Val	Arg	Thr	Met	Gly	Gly	Val	Ser	Ser	Met	Leu	Pro	
	370					375					380					
Lys	Gly	Gly	Asp	Val	Ala	Pro	Asp	Asp	Leu	Asn	Gln	Thr	Asn	Phe	Ser	
385					390					395					400	
Asn	Gly	Ala	Ile	Ile	Arg	Tyr	Arg	Glu	Asp	Ile	Asp	Lys	Asp	His	Asp	
				405					410					415		
Glu	Phe	Asp	Ile	Asp	Asp	Ala	Leu	Gln	Phe	Leu	Lys	Asn	Val	Thr	Asp	
			420					425					430			
Asp	Asp	Phe	Lys	Val	Met	Leu	Ala	Lys	Asn	Tyr	Ser	His	Gly	Leu	Ala	
		435					440					445				
Trp	Thr	Glu	Lys	Glu	Val	Leu	Lys	Asn	Asn	Glu	Met	Pro	Ser	Lys	Trp	
	450					455					460					
Ile	Asn	Pro	Leu	Glu	Thr	Ser	Leu	Pro	Tyr	Ala	Pro	Asp	Met	Lys	Ile	
465					470					475					480	
Tyr	Cys	Val	His	Gly	Val	Gly	Lys	Pro	Thr	Glu	Arg	Gly	Tyr	Tyr	Tyr	
				485					490					495		
Thr	Asn	Asn	Pro	Glu	Gly	Gln	Pro	Val	Ile	Asp	Ser	Ser	Val	Asn	Asp	
			500					505					510			
Gly	Thr	Lys	Val	Glu	Asn	Gly	Ile	Val	Met	Asp	Asp	Gly	Asp	Gly	Thr	
		515					520					525				
Leu	Pro	Ile	Leu	Ala	Leu	Gly	Leu	Val	Cys	Asn	Lys	Val	Trp	Gln	Thr	
	530					535					540					
Lys	Arg	Phe	Asn	Pro	Ala	Asn	Thr	Ser	Ile	Thr	Asn	Tyr	Glu	Ile	Lys	
545					550					555					560	
His	Glu	Pro	Ala	Ala	Phe	Asp	Leu	Arg	Gly	Gly	Pro	Arg	Ser	Ala	Glu	
				565					570					575		
His	Val	Asp	Ile	Leu	Gly	His	Ser	Glu	Leu	Asn	Glu	Ile	Ile	Leu	Lys	
			580					585					590			

23

Val Ser Ser Gly His Gly Asp Ser Val Pro Asn Arg Tyr Ile Ser Asp  
595 600 605

Ile Gln Glu Ile Ile Asn Glu Ile Asn Leu Asp Lys Pro Arg Asn  
610 615 620

<210> 14

<211> 432

<212> PRT

<213> Arabidopsis thaliana

<400> 14

Met Lys Lys Ile Ser Ser His Tyr Ser Val Val Ile Ala Ile Leu Val  
1 5 10 15

Val Val Thr Met Thr Ser Met Cys Gln Ala Val Gly Ser Asn Val Tyr  
20 25 30

Pro Leu Ile Leu Val Pro Gly Asn Gly Gly Asn Gln Leu Glu Val Arg  
35 40 45

Leu Asp Arg Glu Tyr Lys Pro Ser Ser Val Trp Cys Ser Ser Trp Leu  
50 55 60

Tyr Pro Ile His Lys Lys Ser Gly Gly Trp Phe Arg Leu Trp Phe Asp  
65 70 75 80

Ala Ala Val Leu Leu Ser Pro Phe Thr Arg Cys Phe Ser Asp Arg Met  
85 90 95

Met Leu Tyr Tyr Asp Pro Asp Leu Asp Asp Tyr Gln Asn Ala Pro Gly  
100 105 110

Val Gln Thr Arg Val Pro His Phe Gly Ser Thr Lys Ser Leu Leu Tyr  
115 120 125

Leu Asp Pro Arg Leu Arg Asp Ala Thr Ser Tyr Met Glu His Leu Val  
130 135 140

Lys Ala Leu Glu Lys Lys Cys Gly Tyr Val Asn Asp Gln Thr Ile Leu  
145 150 155 160

Gly Ala Pro Tyr Asp Phe Arg Tyr Gly Leu Ala Ala Ser Gly His Pro  
165 170 175

Ser Arg Val Ala Ser Gln Phe Leu Gln Asp Leu Lys Gln Leu Val Glu  
180 185 190

Lys Thr Ser Ser Glu Asn Glu Gly Lys Pro Val Ile Leu Leu Ser His  
195 200 205

Ser Leu Gly Gly Leu Phe Val Leu His Phe Leu Asn Arg Thr Thr Pro  
210 215 220

Ser Trp Arg Arg Lys Tyr Ile Lys His Phe Val Ala Leu Ala Ala Pro  
225 230 235 240

Trp	Gly	Gly	Thr	Ile 245	Ser	Gln	Met	Lys	Thr 250	Phe	Ala	Ser	Gly	Asn 255	Thr
Leu	Gly	Val	Pro 260	Leu	Val	Asn	Pro	Leu 265	Leu	Val	Arg	Arg	His 270	Gln	Arg
Thr	Ser	Glu 275	Ser	Asn	Gln	Trp	Leu 280	Leu	Pro	Ser	Thr	Lys 285	Val	Phe	His
Asp	Arg 290	Thr	Lys	Pro	Leu	Val 295	Val	Thr	Pro	Gln	Val 300	Asn	Tyr	Thr	Ala
Tyr 305	Glu	Met	Asp	Arg	Phe 310	Phe	Ala	Asp	Ile	Gly 315	Phe	Ser	Gln	Gly	Val 320
Val	Pro	Tyr	Lys	Thr 325	Arg	Val	Leu	Pro	Leu 330	Thr	Glu	Glu	Leu	Met 335	Thr
Pro	Gly	Val	Pro 340	Val	Thr	Cys	Ile	Tyr 345	Gly	Arg	Gly	Val	Asp 350	Thr	Pro
Glu	Val	Leu 355	Met	Tyr	Gly	Lys	Gly 360	Gly	Phe	Asp	Lys	Gln 365	Pro	Glu	Ile
Lys	Tyr 370	Gly	Asp	Gly	Asp	Gly 375	Thr	Val	Asn	Leu	Ala 380	Ser	Leu	Ala	Ala
Leu 385	Lys	Val	Asp	Ser	Leu 390	Asn	Thr	Val	Glu	Ile 395	Asp	Gly	Val	Ser	His 400
Thr	Ser	Ile	Leu	Lys 405	Asp	Glu	Ile	Ala	Leu 410	Lys	Glu	Ile	Met	Lys 415	Gln
Ile	Ser	Ile	Ile 420	Asn	Tyr	Glu	Leu	Ala 425	Asn	Val	Asn	Ala	Val	Asn	Glu

```
<210> 15
<211> 552
<212> PRT
<213> Arabidopsis thaliana
```

<400> 15

Met 1	Gly	Ala	Asn	Ser 5	Lys	Ser	Val	Thr	Ala 10	Ser	Phe	Thr	Val	Ile 15	Ala
Val	Phe	Phe	Leu 20	Ile	Cys	Gly	Gly	Arg 25	Thr	Ala	Val	Glu	Asp 30	Glu	Thr
Glu	Phe	His 35	Gly	Asp	Tyr	Ser	Lys 40	Leu	Ser	Gly	Ile	Ile 45	Ile	Pro	Gly
Phe	Ala 50	Ser	Thr	Gln	Leu	Arg 55	Ala	Trp	Ser	Ile	Leu 60	Asp	Cys	Pro	Tyr
Thr 65	Pro	Leu	Asp	Phe	Asn 70	Pro	Leu	Asp	Leu	Val 75	Trp	Leu	Asp	Thr	Thr 80





Met	Gly	Thr	Leu	Phe	Arg	Arg	Asn	Val	Gln	Asn	Gln	Lys	Ser	Asp	Ser	
1				5					10					15		
Asp	Glu	Asn	Asn	Lys	Gly	Gly	Ser	Val	His	Asn	Lys	Arg	Glu	Ser	Arg	
			20					25					30			
Asn	His	Ile	His	His	Gln	Gln	Gly	Leu	Gly	His	Lys	Arg	Arg	Arg	Gly	
		35					40					45				
Ile	Ser	Gly	Ser	Ala	Lys	Arg	Asn	Glu	Arg	Gly	Lys	Asp	Phe	Asp	Arg	
	50					55					60					
Lys	Arg	Asp	Gly	Asn	Gly	Arg	Lys	Arg	Trp	Arg	Asp	Ser	Arg	Arg	Leu	
65					70					75					80	
Ile	Phe	Ile	Leu	Gly	Ala	Phe	Leu	Gly	Val	Leu	Leu	Pro	Phe	Ser	Phe	
				85					90					95		
Gly	Ala	Tyr	His	Val	His	Asn	Ser	Asp	Ser	Asp	Leu	Phe	Asp	Asn	Phe	
			100					105					110			

Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val  
115 120 125

Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn  
130 135 140

Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly  
145 150 155 160

Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val  
165 170 175

Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile  
180 185 190

Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp  
195 200 205

Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp  
210 215 220

Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn  
225 230 235 240

Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile  
245 250 255

Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile  
260 265 270

Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu  
275 280 285

Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys  
290 295 300

Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu  
305 310 315 320

Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp  
325 330 335

Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val  
340 345 350

Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly  
355 360 365

Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr  
370 375 380

Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser  
385 390 395 400

Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser  
405 410 415

Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser

			420				425				430				
Ser	Glu	Asp	Ala	Leu	Asn	Asn	Asn	Thr	Asp	Thr	Tyr	Gly	Asn	Phe	Ile
		435					440					445			
Arg	Phe	Glu	Arg	Asn	Thr	Ser	Asp	Ala	Phe	Asn	Lys	Asn	Leu	Thr	Met
		450					455					460			
Lys	Asp	Ala	Ile	Asn	Met	Thr	Leu	Ser	Ile	Ser	Pro	Glu	Trp	Leu	Gln
465					470					475			480		
Arg	Arg	Val	His	Glu	Gln	Tyr	Ser	Phe	Gly	Tyr	Ser	Lys	Asn	Glu	Glu
				485					490					495	
Glu	Leu	Arg	Lys	Asn	Glu	Leu	His	His	Lys	His	Trp	Ser	Asn	Pro	Met
		500						505					510		
Glu	Val	Pro	Leu	Pro	Glu	Ala	Pro	His	Met	Lys	Ile	Tyr	Cys	Ile	Tyr
		515					520					525			
Gly	Val	Asn	Asn	Pro	Thr	Glu	Arg	Ala	Tyr	Val	Tyr	Lys	Glu	Glu	Asp
		530					535					540			
Asp	Ser	Ser	Ala	Leu	Asn	Leu	Thr	Ile	Asp	Tyr	Glu	Ser	Lys	Gln	Pro
545					550					555			560		
Val	Phe	Leu	Thr	Glu	Gly	Asp	Gly	Thr	Val	Pro	Leu	Val	Ala	His	Ser
				565					570					575	
Met	Cys	His	Lys	Trp	Ala	Gln	Gly	Ala	Ser	Pro	Tyr	Asn	Pro	Ala	Gly
		580						585					590		
Ile	Asn	Val	Thr	Ile	Val	Glu	Met	Lys	His	Gln	Pro	Asp	Arg	Phe	Asp
		595					600					605			
Ile	Arg	Gly	Gly	Ala	Lys	Ser	Ala	Glu	His	Val	Asp	Ile	Leu	Gly	Ser
		610					615					620			
Ala	Glu	Leu	Asn	Asp	Tyr	Ile	Leu	Lys	Ile	Ala	Ser	Gly	Asn	Gly	Asp
625					630					635			640		
Leu	Val	Glu	Pro	Arg	Gln	Leu	Ser	Asn	Leu	Ser	Gln	Trp	Val	Ser	Gln
				645					650					655	
Met	Pro	Phe	Pro	Met											
		660													

```
<210> 17
<211> 387
<212> PRT
<213> Arabidopsis thaliana
```

<400> 17

Val Gly Ser Asn Val Tyr Pro Leu Ile Leu Val Pro Gly Asn Gly Gly  
1 5 10 15

Asn Gln Leu Glu Val Arg Leu Asp Arg Glu Tyr Lys Pro Ser Ser Val



Leu 1	Lys	Lys	Glu	Gly 5	Leu	Lys	Ala	Lys	His 10	Pro	Val	Val	Phe	Ile 15	Pro
Gly	Ile	Val	Thr 20	Gly	Gly	Leu	Glu	Leu 25	Trp	Glu	Gly	Lys	Gln 30	Cys	Ala
Asp	Gly	Leu 35	Phe	Arg	Lys	Arg	Leu 40	Trp	Gly	Gly	Thr	Phe 45	Leu	Cys	Trp
Val	Glu 50	His	Met	Ser	Leu	Asp 55	Asn	Glu	Thr	Gly	Leu 60	Asp	Pro	Ala	Gly
Ile 65	Arg	Val	Arg	Ala	Val 70	Ser	Gly	Leu	Val	Ala 75	Ala	Asp	Tyr	Phe	Ala 80
Pro	Gly	Tyr	Phe	Val 85	Trp	Ala	Val	Leu	Ile 90	Ala	Asn	Leu	Ala	His 95	Ile
Gly	Tyr	Glu	Glu 100	Lys	Asn	Met	Tyr	Met 105	Ala	Ala	Tyr	Asp	Trp 110	Arg	Leu
Ser	Phe	Gln 115	Asn	Thr	Glu	Arg	Asp 120	Gln	Thr	Leu	Ser	Arg 125	Met	Lys	Ser
Asn 130	Ile	Glu	Leu	Met	Val	Ser 135	Thr	Asn	Gly	Gly	Lys 140	Lys	Ala	Val	Ile
Val 145	Pro	His	Ser	Met	Gly 150	Val	Leu	Tyr	Phe	Leu 155	His	Phe	Met	Lys	Trp 160
Val	Glu	Ala	Pro 165	Ala	Pro	Leu	Gly	Gly	Gly 170	Gly	Gly	Pro	Asp	Trp 175	Cys
Ala	Lys	Tyr	Ile 180	Lys	Ala	Val	Met	Asn 185	Ile	Gly	Gly	Pro	Phe 190	Leu	Gly
Val	Pro	Lys 195	Ala	Val	Ala	Gly	Leu 200	Phe	Ser	Ala	Glu	Ala 205	Lys	Asp	Met

20	25	30	
aac cac att cat cat caa cag gga tta ggc cat aag aga aga agg ggt Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly			144
35	40	45	
att agt ggc agt gca aaa aga aat gag cgt ggc aaa gat ttc gac agg Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg			192
50	55	60	
aaa aga gac ggg aac ggt aga aaa cgt tgg aga gat tcc aga aga ctg Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu			240
65	70	75	80
att ttc att ctt ggt gca ttc tta ggt gta ctt ttg ccg ttt agc ttt Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe			288
85	90	95	
ggc gct tat cat gtt cat aat agc gat agc gac ttg ttt gac aac ttt Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe			336
100	105	110	
gta aat ttt gat tca ctt aaa gtg tat ttg gat gat tgg aaa gat gtt Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val			384
115	120	125	
ctc cca caa ggt ata agt tcg ttt att gat gat att cag gct ggt aac Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn			432
130	135	140	
tac tcc aca tct tct tta gat gat ctc agt gaa aat ttt gcc gtt ggt Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly			480
145	150	155	160
aaa caa ctc tta cgt gat tat aat atc gag gcc aaa cat cct gtt gta Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val			528
165	170	175	
atg gtt cct ggt gtc att tct acg gga att gaa agc tgg gga gtt att Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile			576
180	185	190	
gga gac gat gag tgc gat agt tct gcg cat ttt cgt aaa cgg ctg tgg Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp			624
195	200	205	
gga agt ttt tac atg ctg aga aca atg gtt atg gat aaa gtt tgt tgg Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp			672
210	215	220	
ttg aaa cat gta atg tta gat cct gaa aca ggt ctg gac cca ccg aac Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn			720
225	230	235	240
ttt acg cta cgt gca gca cag ggc ttc gaa tca act gat tat ttc atc Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile			768
245	250	255	







Met	Gly	Thr	Leu	Phe	Arg	Arg	Asn	Val	Gln	Asn	Gln	Lys	Ser	Asp	Ser
1				5					10					15	
Asp	Glu	Asn	Asn	Lys	Gly	Gly	Ser	Val	His	Asn	Lys	Arg	Glu	Ser	Arg
			20					25					30		
Asn	His	Ile	His	His	Gln	Gln	Gly	Leu	Gly	His	Lys	Arg	Arg	Arg	Gly
		35					40					45			
Ile	Ser	Gly	Ser	Ala	Lys	Arg	Asn	Glu	Arg	Gly	Lys	Asp	Phe	Asp	Arg
	50					55					60				

Lys 65	Arg	Asp	Gly	Asn	Gly 70	Arg	Lys	Arg	Trp	Arg 75	Asp	Ser	Arg	Arg	Leu 80
Ile	Phe	Ile	Leu	Gly 85	Ala	Phe	Leu	Gly	Val 90	Leu	Leu	Pro	Phe	Ser 95	Phe
Gly	Ala	Tyr	His 100	Val	His	Asn	Ser	Asp 105	Ser	Asp	Leu	Phe	Asp 110	Asn	Phe
Val	Asn	Phe 115	Asp	Ser	Leu	Lys	Val 120	Tyr	Leu	Asp	Asp	Trp 125	Lys	Asp	Val
Leu	Pro 130	Gln	Gly	Ile	Ser	Ser 135	Phe	Ile	Asp	Asp	Ile 140	Gln	Ala	Gly	Asn
Tyr 145	Ser	Thr	Ser	Ser	Leu 150	Asp	Asp	Leu	Ser	Glu 155	Asn	Phe	Ala	Val	Gly 160
Lys	Gln	Leu	Leu	Arg 165	Asp	Tyr	Asn	Ile	Glu 170	Ala	Lys	His	Pro	Val 175	Val
Met	Val	Pro	Gly 180	Val	Ile	Ser	Thr	Gly 185	Ile	Glu	Ser	Trp	Gly 190	Val	Ile
Gly	Asp	Asp 195	Glu	Cys	Asp	Ser	Ser 200	Ala	His	Phe	Arg	Lys 205	Arg	Leu	Trp
Gly	Ser 210	Phe	Tyr	Met	Leu	Arg 215	Thr	Met	Val	Met	Asp 220	Lys	Val	Cys	Trp
Leu 225	Lys	His	Val	Met	Leu 230	Asp	Pro	Glu	Thr	Gly 235	Leu	Asp	Pro	Pro	Asn 240
Phe	Thr	Leu	Arg	Ala 245	Ala	Gln	Gly	Phe	Glu 250	Ser	Thr	Asp	Tyr	Phe 255	Ile
Ala	Gly	Tyr	Trp 260	Ile	Trp	Asn	Lys	Val	Phe	Gln	Asn	Leu	Gly 270	Val	Ile
Gly	Tyr	Glu 275	Pro	Asn	Lys	Met	Thr 280	Ser	Ala	Ala	Tyr	Asp 285	Trp	Arg	Leu
Ala	Tyr 290	Leu	Asp	Leu	Glu	Arg 295	Arg	Asp	Arg	Tyr	Phe 300	Thr	Lys	Leu	Lys
Glu 305	Gln	Ile	Glu	Leu	Phe 310	His	Gln	Leu	Ser	Gly 315	Glu	Lys	Val	Cys	Leu 320
Ile	Gly	His	Ser	Met 325	Gly	Ser	Gln	Ile	Ile 330	Phe	Tyr	Phe	Met	Lys 335	Trp
Val	Glu	Ala	Glu 340	Gly	Pro	Leu	Tyr	Gly 345	Asn	Gly	Gly	Arg	Gly 350	Trp	Val
Asn	Glu	His 355	Ile	Asp	Ser	Phe	Ile 360	Asn	Ala	Ala	Gly	Thr 365	Leu	Leu	Gly
Ala	Pro 370	Lys	Ala	Val	Pro	Ala 375	Leu	Ile	Ser	Gly	Glu 380	Met	Lys	Asp	Thr

<210>	21
<211>	1986

37

atg gtt cct ggt gtc att tct acg gga att gaa agc tgg gga gtt att	576
Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile	
180 185 190	
gga gac gat gag tgc gat agt tct gcg cat ttt cgt aaa cgg ctg tgg	624
Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp	
195 200 205	
gga agt ttt tac atg ctg aga aca atg gtt atg gat aaa gtt tgt tgg	672
Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp	
210 215 220	
ttg aaa cat gta atg tta gat cct gaa aca ggt ctg gac cca ccg aac	720
Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn	
225 230 235 240	
ttt acg cta cgt gca gca cag ggc ttc gaa tca act gat tat ttc atc	768
Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile	
245 250 255	
gca ggg tat tgg att tgg aac aaa gtt ttc caa aat ctg gga gta att	816
Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile	
260 265 270	
ggc tat gaa ccc aat aaa atg acg agt gct gcg tat gat tgg agg ctt	864
Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu	
275 280 285	
gca tat tta gat cta gaa aga cgc gat agg tac ttt acg aag cta aag	912
Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys	
290 295 300	
gaa caa atc gaa ctg ttt cat caa ttg agt ggt gaa aaa gtt tgt tta	960
Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu	
305 310 315 320	
att gga cat tct atg ggt tct cag att atc ttt tac ttt atg aaa tgg	1008
Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp	
325 330 335	
gtc gag gct gaa ggc cct ctt tac ggt aat ggt ggt cgt ggc tgg gtt	1056
Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val	
340 345 350	
aac gaa cac ata gat tca ttc att aat gca gca ggg acg ctt ctg ggc	1104
Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly	
355 360 365	
gct cca aag gca gtt cca gct cta att agt ggt gaa atg aaa gat acc	1152
Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr	
370 375 380	
att caa tta aat acg tta gcc atg tat ggt ttg gaa aag ttc ttc tca	1200
Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser	
385 390 395 400	
aga att gag aga gta aaa atg tta caa acg tgg ggt ggt ata cca tca	1248
Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser	
405 410 415	

Gly	Asp	Asp	Glu	Cys	Asp	Ser	Ser	Ala	His	Phe	Arg	Lys	Arg	Leu	Trp		
			195					200					205				
gga	agt	ttt	tac	atg	ctg	aga	aca	atg	gtt	atg	gat	aaa	gtt	tgt	tgg	672	
Gly	Ser	Phe	Tyr	Met	Leu	Arg	Thr	Met	Val	Met	Asp	Lys	Val	Cys	Trp		
		210					215				220						
ttg	aaa	cat	gta	atg	tta	gat	cct	gaa	aca	ggg	ctg	gac	cca	ccg	aac	720	
Leu	Lys	His	Val	Met	Leu	Asp	Pro	Glu	Thr	Gly	Leu	Asp	Pro	Pro	Asn		240
	225					230					235						
ttt	acg	cta	cgt	gca	gca	cag	ggc	ttc	gaa	tca	act	gat	tat	ttc	atc	768	
Phe	Thr	Leu	Arg	Ala	Ala	Gln	Gly	Phe	Glu	Ser	Thr	Asp	Tyr	Phe	Ile		
				245					250					255			
gca	ggg	tat	tgg	att	tgg	aac	aaa	gtt	ttc	caa	aat	ctg	gga	gta	att	816	
Ala	Gly	Tyr	Trp	Ile	Trp	Asn	Lys	Val	Phe	Gln	Asn	Leu	Gly	Val	Ile		
			260					265						270			
ggc	tat	gaa	ccc	aat	aaa	atg	acg	agt	gct	gcg	tat	gat	tgg	agg	ctt	864	
Gly	Tyr	Glu	Pro	Asn	Lys	Met	Thr	Ser	Ala	Ala	Tyr	Asp	Trp	Arg	Leu		
		275				280						285					
gca	tat	tta	gat	cta	gaa	aga	cgc	gat	agg	tac	ttt	acg	aag	cta	aag	912	
Ala	Tyr	Leu	Asp	Leu	Glu	Arg	Arg	Asp	Arg	Tyr	Phe	Thr	Lys	Leu	Lys		
		290				295					300						
gaa	caa	atc	gaa	ctg	ttt	cat	caa	ttg	agt	ggg	gaa	aaa	gtt	tgt	tta	960	
Glu	Gln	Ile	Glu	Leu	Phe	His	Gln	Leu	Ser	Gly	Glu	Lys	Val	Cys	Leu		
	305				310					315					320		
att	gga	cat	tct	atg	ggg	tct	cag	att	atc	ttt	tac	ttt	atg	aaa	tgg	1008	
Ile	Gly	His	Ser	Met	Gly	Ser	Gln	Ile	Ile	Phe	Tyr	Phe	Met	Lys	Trp		
			325					330						335			
gtc	gag	gct	gaa	ggc	cct	ctt	tac	ggg	aat	ggg	ggg	cgt	ggc	tgg	gtt	1056	
Val	Glu	Ala	Glu	Gly	Pro	Leu	Tyr	Gly	Asn	Gly	Gly	Arg	Gly	Trp	Val		
			340					345					350				
aac	gaa	cac	ata	gat	tca	ttc	att	aat	gca	gca	ggg	acg	ctt	ctg	ggc	1104	
Asn	Glu	His	Ile	Asp	Ser	Phe	Ile	Asn	Ala	Ala	Gly	Thr	Leu	Leu	Gly		
		355				360						365					
gct	cca	aag	gca	gtt	cca	gct	cta	att	agt	ggg	gaa	atg	aaa	gat	acc	1152	
Ala	Pro	Lys	Ala	Val	Pro	Ala	Leu	Ile	Ser	Gly	Glu	Met	Lys	Asp	Thr		
		370				375					380						
att	caa	tta	aat	acg	tta	gcc	atg	tat	ggg	ttg	gaa	aag	ttc	ttc	tca	1200	
Ile	Gln	Leu	Asn	Thr	Leu	Ala	Met	Tyr	Gly	Leu	Glu	Lys	Phe	Phe	Ser		
		385			390					395					400		
aga	att	gag	aga	gta	aaa	atg	tta	caa	acg	tgg	ggg	ggg	ata	cca	tca	1248	
Arg	Ile	Glu	Arg	Val	Lys	Met	Leu	Gln	Thr	Trp	Gly	Gly	Ile	Pro	Ser		
			405					410						415			
atg	cta	cca	aag	gga	gaa	gag	gtc	att	tgg	ggg	gat	atg	aag	tca	tct	1296	
Met	Leu	Pro	Lys	Gly	Glu	Glu	Val	Ile	Trp	Gly	Asp	Met	Lys	Ser	Ser		
			420					425					430				

tca gag gat gca ttg aat aac aac act gac aca tac ggc aat ttc att	1344
Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile	
435 440 445	
cga ttt gaa agg aat acg agc gat gct ttc aac aaa aat ttg aca atg	1392
Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met	
450 455 460	
aaa gac gcc att aac atg aca tta tcg ata tca cct gaa tgg ctc caa	1440
Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln	
465 470 475 480	
aga aga gta cat gag cag tac tcg ttc ggc tat tcc aag aat gaa gaa	1488
Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu	
485 490 495	
gag tta aga aaa aat gag cta cac cac aag cac tgg tcg aat cca atg	1536
Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met	
500 505 510	
gaa gta cca ctt cca gaa gct ccc cac atg aaa atc tat tgt ata tac	1584
Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr	
515 520 525	
ggg gtg aac aac cca act gaa agg gca tat gta tat aag gaa gag gat	1632
Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp	
530 535 540	
gac tcc tct gct ctg aat ttg acc atc gac tac gaa agc aag caa cct	1680
Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro	
545 550 555 560	
gta ttc ctc acc gag ggg gac gga acc gtt ccg ctc gtg gcg cat tca	1728
Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser	
565 570 575	
atg tgt cac aaa tgg gcc cag ggt gct tca ccg tac aac cct gcc gga	1776
Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly	
580 585 590	
att aac gtt act att gtg gaa atg aaa cac cag cca gat cga ttt gat	1824
Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp	
595 600 605	
ata cgt ggt gga gca aaa agc gcc gaa cac gta gac atc ctc ggc agc	1872
Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser	
610 615 620	
gcg gag ttg aac gat tac atc ttg aaa att gca agc ggt aat ggc gat	1920
Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp	
625 630 635 640	
ctc gtc gag cca cgc caa ttg tct aat ttg agc cag tgg gtt tct cag	1968
Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln	
645 650 655	
atg ccc ttc cca atg taa	1986
Met Pro Phe Pro Met	
660	

<210> 22  
 <211> 661  
 <212> PRT  
 <213> *Saccharomyces cerevisiae*

<400> 22

Met	Gly	Thr	Leu	Phe	Arg	Arg	Asn	Val	Gln	Asn	Gln	Lys	Ser	Asp	Ser	1	5	10	15
Asp	Glu	Asn	Asn	Lys	Gly	Gly	Ser	Val	His	Asn	Lys	Arg	Glu	Ser	Arg	20	25	30	
Asn	His	Ile	His	His	Gln	Gln	Gly	Leu	Gly	His	Lys	Arg	Arg	Arg	Gly	35	40	45	
Ile	Ser	Gly	Ser	Ala	Lys	Arg	Asn	Glu	Arg	Gly	Lys	Asp	Phe	Asp	Arg	50	55	60	
Lys	Arg	Asp	Gly	Asn	Gly	Arg	Lys	Arg	Trp	Arg	Asp	Ser	Arg	Arg	Leu	65	70	75	80
Ile	Phe	Ile	Leu	Gly	Ala	Phe	Leu	Gly	Val	Leu	Leu	Pro	Phe	Ser	Phe	85	90	95	
Gly	Ala	Tyr	His	Val	His	Asn	Ser	Asp	Ser	Asp	Leu	Phe	Asp	Asn	Phe	100	105	110	
Val	Asn	Phe	Asp	Ser	Leu	Lys	Val	Tyr	Leu	Asp	Asp	Trp	Lys	Asp	Val	115	120	125	
Leu	Pro	Gln	Gly	Ile	Ser	Ser	Phe	Ile	Asp	Asp	Ile	Gln	Ala	Gly	Asn	130	135	140	
Tyr	Ser	Thr	Ser	Ser	Leu	Asp	Asp	Leu	Ser	Glu	Asn	Phe	Ala	Val	Gly	145	150	155	160
Lys	Gln	Leu	Leu	Arg	Asp	Tyr	Asn	Ile	Glu	Ala	Lys	His	Pro	Val	Val	165	170	175	
Met	Val	Pro	Gly	Val	Ile	Ser	Thr	Gly	Ile	Glu	Ser	Trp	Gly	Val	Ile	180	185	190	
Gly	Asp	Asp	Glu	Cys	Asp	Ser	Ser	Ala	His	Phe	Arg	Lys	Arg	Leu	Trp	195	200	205	
Gly	Ser	Phe	Tyr	Met	Leu	Arg	Thr	Met	Val	Met	Asp	Lys	Val	Cys	Trp	210	215	220	
Leu	Lys	His	Val	Met	Leu	Asp	Pro	Glu	Thr	Gly	Leu	Asp	Pro	Pro	Asn	225	230	235	240
Phe	Thr	Leu	Arg	Ala	Ala	Gln	Gly	Phe	Glu	Ser	Thr	Asp	Tyr	Phe	Ile	245	250	255	
Ala	Gly	Tyr	Trp	Ile	Trp	Asn	Lys	Val	Phe	Gln	Asn	Leu	Gly	Val	Ile	260	265	270	



Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu  
 275 280 285  
 Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys  
 290 295 300  
 Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu  
 305 310 315 320  
 Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp  
 325 330 335  
 Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val  
 340 345 350  
 Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly  
 355 360 365  
 Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr  
 370 375 380  
 Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser  
 385 390 395 400  
 Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser  
 405 410 415  
 Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser  
 420 425 430  
 Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile  
 435 440 445  
 Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met  
 450 455 460  
 Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln  
 465 470 475 480  
 Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu  
 485 490 495  
 Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met  
 500 505 510  
 Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr  
 515 520 525  
 Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp  
 530 535 540  
 Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro  
 545 550 555 560  
 Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser  
 565 570 575  
 Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly

580 585 590

Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp  
595 600 605

Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser  
610 615 620

Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp  
625 630 635 640

Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln  
645 650 655

Met Pro Phe Pro Met  
660

<210> 23  
<211> 2312  
<212> DNA  
<213> Schizosaccharomyces pombe

<400> 23

atggcgctctt ccaagaagag caaaactcat aagaaaaaga aagaagtcaa atctcctatc 60  
gacttaccaa attcaaagaa accaactcgc gctttgagtg agcaaccttc agcgtccgaa 120  
acacaatctg tttcaaataa atcaagaaaa tctaaatttg gaaaaagatt gaattttata 180  
ttgggcgcta ttttgggaat atgcggtgct ttttttttcg ctgttggaga cgacaatgct 240  
gttttcgacc ctgctacgtt agataaattt gggaatatgc taggctcttc agacttgttt 300  
gatgacatta aaggatattt atcttataat gtgtttaagg atgcaccttt tactacggac 360  
aagccttcgc agtctcctag cggaaatgaa gttcaagttg gtcttgatat gtacaatgag 420  
ggatatcgaa gtgaccatcc tgttattatg gttcctggtg ttatcagctc aggattagaa 480  
agttggtcgt ttaataattg ctcgattcct tacttttagga aacgtctttg gggtagctgg 540  
tctatgctga aggcaatggt ccttgacaag caatgctggc ttgaacattt aatgcttgat 600  
aaaaaaaccg gcttgatcc gaagggaatt aagctgcgag cagctcaggg gtttgaagca 660  
gctgattttt ttatcacggg ctattggatt tggagtaaag taattgaaaa ccttgctgca 720  
attggttatg agcctaataa catgttaagt gcttcctacg attggcggtt atcatatgca 780  
aatntagagg aacgtgataa atatttttca aagttaaaaa tggttcattga gtacagcaac 840  
attgtacata agaaaaaggt agtggtgatt tctcactcca tgggttcaca gggttacgtac 900  
tattttttta agtgggttga agctgagggc tacggaaatg gtggaccgac ttgggttaat 960  
gatcatattg aagcatttat aaatgtgagt ctcgatgggt gtttgactac gtttctaact 1020

43

```

tttgaataga tatcgggatac tttgattgga gcacccaaaa cagtggcagc gcttttatcg 1080
ggtgaaatga aagatacagg tattgtaatt acattaaaca tggttaatatt taattttttgc 1140
taaccgtttt aagctcaatt gaatcagttt tcgggtctatg ggtaagcaat aaattggtga 1200
gatttggttac taatttactg tttagtttgg aaaaattttt ttcccgttct gaggtatatt 1260
caaaaataca aatgtgctct actttttcta acttttaata gagagccatg atgggttcgca 1320
ctatgggagg agttagtctt atgcttccta aaggaggcga tgttgatgg ggaaatgccca 1380
gttgggtaag aaatatgtgc tgttaatttt ttattaatat ttaggctcca gatgatctta 1440
atcaaacaaa tttttccaat ggtgcaatta ttogatatag agaagacatt gataaggacc 1500
acgatgaatt tgacatagat gatgcattac aatttttaaa aaatgttaca gatgacgatt 1560
ttaaagtcac gctagcgaaa aattattccc acggtcttgc ttggactgaa aaagaagtgt 1620
taaaaaataa cgaaatgccg tctaaatgga taaatccgct agaagtaaga acattaaagt 1680
tactaaatta tactaaccga aatagactag tcttcttat gctcctgata tgaaaattta 1740
ttgcgttcac ggggtcggaa aaccaactga gagagggttat tattatacta ataatcctga 1800
ggggcaacct gtcattgatt cctcgggtta tgatggaaca aaagttgaaa atgtgagaga 1860
atztatgttt caaacattct attaaactgt ttattagggt attggttatgg atgatgggtga 1920
tggaacttta ccaatattag cccttggttt ggtgtgcaat aaagtttggc aaacaaaaag 1980
gtttaatcct gctaatacaa gtatcacaaa ttatgaaatc aagcatgaac ctgctgcgtt 2040
tgatctgaga ggaggacctc gctcggcaga acacgtcgat atacttggaac attcagagct 2100
aaatgtatgt tcattttacc ttacaaattt ctattactaa ctcttgaaat aaggaaatta 2160
ttttaaaagt ttcattcaggc catggtgact cggtaacaaa ccgttatata tcagatatcc 2220
agtacggaca taagttttgt agattgcaat taactaacta accgaacagg gaaataataa 2280
atgagataaa tctcgataaa cctagaaatt aa 2312

```

<210> 24  
 <211> 3685  
 <212> DNA  
 <213> Arabidopsis thaliana

<400> 24

```

atgccccctta ttcattcggaa aaagccgacg gagaaaccat cgacgccgcc atctgaagag 60
gtggtgcacg atgaggattc gcaaaagaaa ccacacgaat cttccaaatc ccaccataag 120
aaatcgaacg gaggagggaa gtggtcgtgc atcgattctt gttggtgggt cattgggtgt 180
gtgtgtgtaa cctgggtggtt tcttctcttc ctttacaacg caatgcctgc gagcttcctt 240

```

cagtatgtaa cggagcgaat cacgggtcct ttgectgacc cgcccgggtg taagctcaaa 300  
aaagaaggtc ttaaggcgaa acatcctggt gtcttcattc ctgggattgt caccgggtggg 360  
ctcgagcttt gggaaggcaa acaatgcgct gatggtttat ttagaaaacg tttgtggggg 420  
ggaacttttg gtgaagtcta caaaagggtga gctcaacaat tctcactctt cctttatatt 480  
gggatttgga ttggatctga tgagatcacg cacttggtgc ttcttcaaca tcaactcaaac 540  
tttaattcca tgtttgtctg tcttactctt taactttttt tttttttgat gtgaaacgct 600  
attttcttaa gagactatct ctgtatgtgt aaggtaagcg ttccaaggac gtaattgggt 660  
tggactatct ctgtttgatt gttaacttta ggatataaaa tagctgcctt ggaatttcaa 720  
gtcatcttat tgccaaatct gttgctagac atgccctaga gtccgttcat aacaagttac 780  
ttcctttact gtcgttgctg gttagattag ctttgtgtag cgtataatga agtagtggtt 840  
tatgttttgt tgggaataga gaagttctaa ctacatctgt ggaaagtgtg ttcaggctgt 900  
gatagaggac tgttgcttta ttattcaact atgtatatgt gtaattaaag ctagtctcct 960  
tttgatcttt cagctcaatg tgcttttctc aatttttttc tcaatttcaa agtttcacat 1020  
cgagtttatt cacatgtctt gaatttcgtc catcctcggt ctgttatcca gctttgaact 1080  
cctcccgacc ctgctatgga tatattaaaa aaaaagtgtt ttgtgggttg catctttgtt 1140  
acgatctgca tcttcttctt tcggctcagt gttcatgttt ttgctatggt agagatgggc 1200  
aatgttattg ttgatggtaa cagtgggtata gttgatagta tcttaactaa tcaattatct 1260  
ctttgattca ggctctatg ttgggtggaa cacatgtcac ttgacaatga aactgggttg 1320  
gatccagctg gtattagagt tcgagctgta tcaggactcg tggctgctga ctactttgct 1380  
cctggctact ttgtctgggc agtgctgatt gctaaccttg cacatattgg atatgaagag 1440  
aaaaatatgt acatggctgc atatgactgg cggttttcgt ttcagaacac agaggttctt 1500  
ttctcatcgt tctttctatt attctgttcc atgttacgtt tctttcttca ttacttaagg 1560  
cttaaatatg tttcatgttg aattaatagg tacgtgatca gactcttagc cgtatgaaaa 1620  
gtaatataga gttgatggtt tctaccaacg gtggaaaaaa agcagttata gttccgcatt 1680  
ccatgggggt cttgtatctt ctacatttta tgaagtgggt tgaggcacca gctcctctgg 1740  
gtggcggggg tgggccagat tgggtgtgcaa agtatattaa ggcgggtgat aacattgggt 1800  
gaccatttct tgggtgtcca aaagctgttg cagggtcttt ctctgctgaa gcaaaggatg 1860  
ttgcagttgc caggtattga atatctgctt atacttttga tgatcagaac cttggctctg 1920  
gaactcaaag ttattctact aaatatcaat tctaataaca ttgctatatt atcgctgcaa 1980  
ctgacattgg ttgattatct ttgctgctta tgtaactgaa actctcttga gattagacaa 2040

atgatgaatt gataattctt acgcattgct ctgtgatgac cagtttctta gcttcgacga 2100  
taacatttgt catactgtct tttggagggc attgaatttt gctatggaaa gcgctggagc 2160  
ttccatgctt gcattcttta ccaattagcg ttattctgct tctttcaatt ttcttgtata 2220  
tgcacttatg gtcttttatt tcttcttaat taaagactcg ttggattagt tgctctatta 2280  
gtcacttggt tccttaatat agaactttac tttcttcgaa aattgcagag cgattgcccc 2340  
aggattctta gacaccgata tathtagact tcagaccttg cagcatgtaa tgagaatgac 2400  
acgcacatgg gactcaacaa tgtctatggt accgaagggg ggtgacacga tatggggcgg 2460  
gcttgattgg tcaccggaga aaggccacac ctgttgtggg aaaaagcaaa agaacaacga 2520  
aacttgtggt gaagcaggtg aaaacggagt ttccaagaaa agtctgttta actatggaag 2580  
gatgatatct tttgggaaag aagtagcaga ggctgcgcca tctgagatta ataatttga 2640  
ttttcgagta aggacatata aatcataata aacctgttac attttgtgat tgtatgatga 2700  
atatctgtac attttatctg gtgaaggggtg ctgtcaaagg tcagagtatc ccaaatacaca 2760  
cctgtcgtga cgtgtggaca gagtaccatg acatgggaat tgctgggacg aaagctatcg 2820  
ctgagtataa ggtctacact gctggtgaag ctatagatct actacattat gttgctccta 2880  
agatgatggc gcgtgggtgc gctcatttct cttatggaat tgctgatgat ttggatgaca 2940  
ccaagtatca agatcccaaa tactgggtcaa atccgttaga gacaaagtaa gtgatttctt 3000  
gattccaact gtatccttcg tcctgatgca ttatcagtct ttttgttttc ggtcttgttg 3060  
gatatggttt tcagctcaaa gcttacaaag ctgtttctga gcctttctca aaaaggcttg 3120  
ctcagtaata ttgaggtgct aaagttgata catgtgactc ttgcttataa atcctccgtt 3180  
tggtttgttc tgctttttca gattaccgaa tgctcctgag atggaaatct actcattata 3240  
cggagtgggg ataccaacgg aacgagcata cgtatacaag cttaccagct ctcccgacag 3300  
ttgcatcccc ttccagatat tcacttctgc tcacgaggag gacgaagata gctgtctgaa 3360  
agcaggagtt tacaatgtgg atggggatga aacagtaccc gtcttaagtg ccgggtacat 3420  
gtgtgcaaaa gcgtggcgtg gcaagacaag attcaaccct tccggaatca agacttatat 3480  
aagagaatac aatcactctc cgccggctaa cctgttgga gggcgcgga cgagagtgg 3540  
tgcccatggt gatatcatgg gaaactttgc tttgatcgaa gatatcatga gggttgccgc 3600  
cggaggtaac ggtctgata taggacatga ccagggtccac tctggcatat ttgaatggtc 3660  
ggagcgtatt gacctgaagc tgtga 3685

<212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (120)..(401)

<220>  
 <221> unsure  
 <222> 1..402  
 <223> n= a or g or c or t/u  
 <223> Xaa=unknown

<400> 25

```

agaacacagct ctttgtctct ctcgactgat ctaacaatcc ctaatctgtg ttctaaattc 60
ctggacgaga tttgacaaag tccgtatagc ttaacctggt ttaatttcaa gtgacagat 119
atg ccc ctt att cat cgg aaa aag ccg acg gag aaa cca tcg acg ccg 167
Met Pro Leu Ile His Arg Lys Lys Pro Thr Glu Lys Pro Ser Thr Pro
  1             5             10             15
cca tct gaa gag gtg gtg cac gat gag gat tcg caa aag aaa cca cac 215
Pro Ser Glu Glu Val Val His Asp Glu Asp Ser Gln Lys Lys Pro His
          20             25             30
gaa tct tcc aaa tcc cac cat aag naa tcg aac gga gga ggg aag tgg 263
Glu Ser Ser Lys Ser His His Lys Xaa Ser Asn Gly Gly Gly Lys Trp
          35             40             45
tcg tgc atc gat tct tgt tgt tgg ttc att ggg tgt gtg tgt gta acc 311
Ser Cys Ile Asp Ser Cys Cys Trp Phe Ile Gly Cys Val Cys Val Thr
          50             55             60
tgg tgg ttt ctt ctc ttc ctt tac aac gca atg cct gcg agc ttc cct 359
Trp Trp Phe Leu Leu Phe Leu Tyr Asn Ala Met Pro Ala Ser Phe Pro
          65             70             75             80
cag tat gta acg gag ccg aat cac gng tcc ttt gcc tta ccc g 402
Gln Tyr Val Thr Glu Pro Asn His Xaa Ser Phe Ala Leu Pro
          85             90

```

<210> 26  
 <211> 643  
 <212> DNA  
 <213> Zea mays

<220>  
 <221> CDS  
 <222> (1)..(402)

<220>  
 <221> unsure  
 <222> 1..643  
 <223> n= a or g or c or t/u

<400> 26

Arg	Glu	Lys	Ile	Ala	Ala	Leu	Lys	Gly	Gly	Val	Tyr	Leu	Ala	Asp	Gly
1				5					10					15	
Asp	Glu	Thr	Val	Pro	Val	Leu	Ser	Ala	Gly	Tyr	Met	Cys	Ala	Lys	Gly
			20					25					30		
Trp	Arg	Gly	Lys	Thr	Arg	Phe	Ser	Pro	Ala	Gly	Ser	Lys	Thr	Tyr	Val
		35					40					45			

48

Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly  
 50 55 60  
 Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile  
 65 70 75 80  
 Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly  
 85 90 95  
 Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Lys  
 100 105 110  
 Leu Lys Leu  
 115

<210> 28  
 <211> 516  
 <212> DNA  
 <213> Neurospora crassa

<220>  
 <221> unsure  
 <222> 1..516  
 <223> n= a or g or c or t/u

<400> 28

ggtggcgaag acganggcgg aagttggagg ctaacgagaa tgacnctcgg agatggatct 60  
 accctctaga gacacgacta ccnttgcacc cagcctcaag gtntacngtt tntatgggta 120  
 ggaagccgac ggagcgagcc tacatctatc tggcgcccgga tcccgggacg acaacgcac 180  
 tttgatgac gatcgatacg actttgactn aggggcacat tgaccacggt gtgatttttg 240  
 gcgaaggcga tggcacagtg aaccttatga gtttggggta cctgtgcaat aaggggtgga 300  
 aaatgaagag atacaatcct gcgggctcaa aaataaccgt ggtcgagatg ccgcatgaac 360  
 cagaacggtt caatccgaga ggagggccga atacggcgga cttaaataatg tagaaaaggt 420  
 tgaaatttat gaagagtaat taaatacggc acatagggtta ctcaatagta tgactaatta 480  
 aaaaaaaatt ttttttctaa aaaaaaaaaa aaaaaa 516

<210> 29  
 <211> 1562  
 <212> DNA  
 <213> Arabidopsis thaliana

<400> 29

atgaaaaaaaa tatcttcaca ttattcggta gtcatacgga tactcgttgt ggtgacgatg 60  
 acctcgatgt gtcaagctgt gggtagcaac gtgtaccctt tgattctggt tccaggaaac 120  
 ggaggttaacc agctagaggt acggctggac agagaatata agccaagtag tgtctggtgt 180



49

```

agcagctggt tatatccgat tcataagaag agtgggtggat ggttttaggct atgggttcgat 240
gcagcagtgt tattgtctcc cttcaccagg tgcttcagcg atcgaatgat gttgtactat 300
gacctgatt tggatgatta ccaaaatgct cctgggtgtcc aaaccgggt tcctcatttc 360
ggttcgacca aatcacttct atacctcgac cctcgtctcc ggtagtact ttccaagata 420
tatcattttg ggacatttgc ataatgaaca aaatagacat aaatttgggg gattattgtt 480
atatcaatat ccatttatat gctagtcggt aatgtgagtg ttatgttagt atagttaatg 540
tgagtgttat gtgattttcc attttaaatg aagctagaaa gttgtcgttt aataatgttg 600
ctatgtcatg agaattataa ggacactatg taaatgtagc ttaataataa ggtttgattt 660
gcagagatgc cacatcttac atggaacatt tggtgaaagc tctagagaaa aaatgctggg 720
atgttaacga ccaaaccatc ctaggagctc catatgattt caggtacggc ctggctgctt 780
cgggccaccc gtcccgtgta gctcacagt tcctacaaga cctcaaaca ttggtggaaa 840
aaactagcag cgagaacgaa ggaaagccag tgatactcct ctcccatagc ctaggaggac 900
tttctgctct ccatttcctc aaccgtacca ccccttcctg gcgcgcgaag tacatcaaac 960
actttgttgc actcgtgctg ccatggggtg ggacgatctc tcagatgaag acatttgctt 1020
ctggcaacac actcgggtgt ccttttagtta accctttgct ggtcagacgg catcagagga 1080
cctccgagag taaccaatgg ctacttccat ctaccaaagt gtttcacgac agaactaaac 1140
cgcttgctgt aactccccag gttaactaca cagcttacga gatggatcgg ttttttgcag 1200
acattggatt ctacaagga gttgtgcctt acaagacaag agtggtgcct ttaacagagg 1260
agctgatgac tccgggagtg ccagtcactt gcatatatgg gagaggagt gatacaccgg 1320
aggttttgat gtatggaaaa ggaggattcg ataagcaacc agagattaag tatggagatg 1380
gagatgggac ggttaatttg gcgagcttag cagctttgaa agtcgatagc ttgaacaccg 1440
tagagattga tggagtttcg catacatcta tacttaaaga cgagatcgca cttaaagaga 1500
ttatgaagca gatttcaatt attaatatg aattagccaa tgttaatgcc gtcaatgaat 1560
ga 1562

```

```

<210> 30
<211> 3896
<212> DNA
<213> Arabidopsis thaliana
<400> 30

```

```

atgggagcga attcgaaatc agtaacggct tccttcaccg tcacgcgcgt ttttttcttg 60
atttgcggtg gccgaactgc ggtggaggat gagaccgagt ttcacggcga ctactcgaag 120

```

ctatcgggta taatcattcc gggatttgcg tcgacgcagc tacgagcgtg gtcgatcctt 180  
gactgtccat acactccggt ggacttcaat ccgctcgacc tcgtatggct agacaccact 240  
aaggtcggtg atcttcattt ccttcgctcc ttattctgtc ggtcgagtc aattgttgatg 300  
aattccaagc gaaatatagc aatgaagcat gtctcgctc tcttattgat tcgttcatta 360  
gtcaacagtg acgcttctga atctgagttt agagtcatat aaaacagctg actcggcgag 420  
tgtttcccat cgcttttggt tcgctaaatg tagcgcaatg aatgtgtaat tagtctgcgc 480  
tttttattca actagatctg caagtttttc agagtgtc aatagtagtta gaaaatgtta 540  
ggtcatttta cttgtgcatt gtgattcttt tggttgttgc ttactgatcg acgtgatgga 600  
tggtttacag cttctttctg ctgtcaactg ctggtttaag tgtatgggtg tagatcctta 660  
taatcaaaca gaccatcccg agtgtaagtc acggcctgac agtgggtctt cagccatcac 720  
agaattggat ccagggttaca taacaggtag ttccggattt ttctttcttt tgagtcttct 780  
tcaatttgat atcatcttgt tgtgatataa tatggctaag ttcattaatt tgggtcaattt 840  
tcaggctctc tttctactgt ctggaaagag tggcttaagt ggtgtgttga gtttggtata 900  
gaagcaaatg caattgtcgc tgttccatac gattggagat tgtcaccaac caaattggaa 960  
gagcgtgacc tttactttca caagctcaag ttagtcctta tcaggcta atgtcttttct 1020  
ttctcttttt atgtaagata agctaagagc tctggctgct ttcctttttg cagggtgacc 1080  
tttgaaactg ctttaaaaact ccgtggcggc ccttctatag tatttgccca ttcaatgggt 1140  
aataatgtct tcagatactt tctggaatgg ctgaggctag aaattgcacc aaaacattat 1200  
ttgaagtggc ttgatcagca tatccatgct tatttcgctg ttggtaccgg cctactatcc 1260  
ttaagttacc attttatttt ttctctaatt gggggaggtta tgttgtgact tactggattg 1320  
agctcgatac ctgatttggt gttgatttag gagctcctct tcttggttct gttgaggcaa 1380  
tcaaactctac tctctctggt gtaacgtttg gccttcctgt ttctgaggtg acctctgact 1440  
tctcttttagt ttttaagtag tgatatcaac caggctttat aactcactgg attttccttt 1500  
tgaaagtatt acttttggtt attgaactgc tgaacgcgat atggtatctg tagatcttga 1560  
agtgttagtt atcaaagaac atattgtggg tagtatacct gtcagcggcc ttagctaata 1620  
caaccaaacc acatgtacac tgatttagtt ttcagattat tatggtagac ttttaagttga 1680  
gaagaaactt tgactgaaat ctttttattt taataggcta tgatttggtt attgaaatca 1740  
tgtgacatat tgacatgcgc ttctcatggt ttttggtggc aaggcttcag ggaactgtc 1800  
ggttgtgtgc caattctttt gcgtcgctat tgtggcttat gccattttca agaattgca 1860  
agggtgataa cacattctgg acgcattttt ctgggggtgc tgcaaagaaa gataagcgcg 1920

tataccactg tgatgaagag gaatatcaat caaaatattc tggctggccg acaaataatta 1980  
 ttaacattga aattccttcc actagcgggt agactctgta tatgcaactg taacactaac 2040  
 aaaagtttca ccaagaatgt tcactctcat atttcgttcc tttgatgtgt atccatcagt 2100  
 tacagaaaca gctctagtca acatgaccag catggaatgt ggccttccca cccttttgtc 2160  
 tttcacagcc cgtgaactag cagatgggac tcttttcaaa gcaatagaag actatgaccc 2220  
 agatagcaag aggatgttac accagttaaa gaagtacgta cctttctttg tgataagaaa 2280  
 tattgctcat cgatcatcac ttgctggctt cttgtacgtc aaattgtttt gtttaaactt 2340  
 ctatatcaat tgttcatatg ctttgtcttt cttactataa gaaacaagta taatcagaaa 2400  
 ccttattatt gattatcagt tctctcctta tattatggaa tgtctttttc gtttacagtt 2460  
 atgaatgcaa aagggggtat tttagtgtat tgattctctc attctctagt ttgttttgac 2520  
 taatagcgtc aattttgttt ttctagcaaa tctttgtgaa ttatatataa catgctaact 2580  
 atacttttca ggttgatca tgatgaccct gtttttaata ctctgactcc ttgggagaga 2640  
 ccacctataa aaaatgtatt ttgcatatat ggtgctcacc taaagacaga ggtatgatgc 2700  
 attctcaata tcacattatg cgttgacttt gttattatat tccccatttg gtttgcaata 2760  
 tctttttgaa ttatgattta tcttctccct tgcactttat gctattaagc gttaaaggta 2820  
 ctaaagtat gaagctgtct gtcatagggt gggtattact ttgccccagc tggcaaacct 2880  
 tatectgata attggatcat cacggatata atttatgaaa ctgaagggtc cctcgtgtca 2940  
 aggtaatttt ccgcaatggc agaagtaaaa caggaaggca aagtcttctg tatcagtcta 3000  
 gtggcatgtt atctcagttg cataagcaaa ttattaaaca actaaaattt aagtactttt 3060  
 ttatcattcc ttttgagctt agtggatgat cagtggctta aagtgggaag aggtgttgca 3120  
 tgaaacatga cacttgatc aaagataact agcaaaacaa aactaaccba tttctgaatt 3180  
 tcatattatt aggagtagtc gtgcttttaa aaaatttggt ttaagaaacc gaaaaactag 3240  
 ttcatatctt gatttgcaaa tatctgcagg tctggaactg tggttgatgg gaacgctgga 3300  
 cctataactg gggatgagac ggtaagctca gaagttgggt ttgaaattat cttcttgcaa 3360  
 actactgaag actaagataa tacttgcttc tggaacactg cttgctatgt tctctagtac 3420  
 actgcaatat tgactctccg ctacttttat tgattatgaa attgatctct tataggtacc 3480  
 ctatcattca ctctcttggg gcaagaattg gctcggacct aaagttaaca taacaatggc 3540  
 tccccaggta ctctttttta gttcctcacc ttatatagat caaactttta gtgtactttt 3600  
 ctggttatgt gttgatttac ctccaatttg ttctttctaa aaatcatata tctctgtact 3660  
 cctcaagaac ttgtattaat ctaaacgaga ttctcattgg gaaaataaaa caacagccag 3720

aacacgatgg aagcgacgta catgtggaac taaatgttga tcatgagcat gggtcagaca 3780  
 tcatagctaa catgacaaaa gcaccaaggg ttaagtacat aaccttttat gaagactctg 3840  
 agagcattcc ggggaagaga accgcagtct gggagcttga taaaagtggg tattaa 3896

<210> 31  
 <211> 709  
 <212> DNA  
 <213> tomato

<400> 31

ctggggccaa aagtgaacat aacaaggaca ccacagtcag agcatgatgt tcagatgtac 60  
 aagtgcattct aaatatagag catcaacatg gtgaagatat cattcccaat atgacaaagt 120  
 tacctacaat gaagtacata acctattatg aggattctga aagttttcca gggacaagaa 180  
 cagcagtttg ggagcttgat aaagcaaata acaggaacat tgtcagatct ccagctttga 240  
 tgcgggagct gtggcttgag atgtggcatg atattcatcc tgataaaaag tccaagtttg 300  
 ttacaaaagg tgggtgtctga tctcactat tttctctat aaatgtttga gtttgtattg 360  
 acattgtaag tattgcaaca aaaagcaaag cgtgggcctc tgagggatga ggactgctat 420  
 tgggattacg ggaaagctcg atgtgcatgg gctgaacatt gtgaatacag gttagaatat 480  
 tcaaattata ttttgcaaaa tattctcttt ttgtgtattht aggccacctt tccccgggtca 540  
 caacgatgca gatatgtatt cggggatgtht cacctgggac agagttgcag attgaagagt 600  
 tctacatctc acatcctgtc aactatgtg tgatatttaa gaaactttgt ttggcggaac 660  
 aacaagtttg cacaacatt tgaagaagaa agcgaaatga ttcagagag 709

09/937779

JC05 Rec'd PCT/PTO 2 8 SEP 2001

## SEQUENCE LISTING

&lt;110&gt; BASF AG

<120> A NEW CLASS OF ENZYMES IN THE BIOSYNTHETIC PATHWAY FOR  
THE PRODUCTION OF TRIACYLGLYCEROL AND RECOMBINANT DNA  
MOLECULES ENCODING THESE ENZYMES

&lt;130&gt; BASF-NAE-3377-99-Sept-2000

&lt;140&gt; PCT/EP 00/02701

&lt;141&gt; 2000-03-23

&lt;160&gt; 31

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 1986

&lt;212&gt; genomic DNA

&lt;213&gt; Saccharomyces cerevisiae

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1983)

&lt;400&gt; 1

```

atg ggc aca ctg ttt cga aga aat gtc cag aac caa aag agt gat tct      48
Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser
   1             5             10             15

gat gaa aac aat aaa ggg ggt tct gtt cat aac aag cga gag agc aga      96
Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg
           20           25           30

aac cac att cat cat caa cag gga tta ggc cat aag aga aga agg ggt      144
Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly
       35           40           45

att agt ggc agt gca aaa aga aat gag cgt ggc aaa gat ttc gac agg      192
Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg
   50           55           60

aaa aga gac ggg aac ggt aga aaa cgt tgg aga gat tcc aga aga ctg      240
Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu
   65           70           75           80

att ttc att ctg ggt gca ttc tta ggt gta ctg ttg ccg ttt agc ttt      288
Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe
       85           90           95

ggc gct tat cat gtt cat aat agc gat agc gac ttg ttt gac aac ttt      336
Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe
      100           105           110

gta aat ttt gat tca ctt aaa gtg tat ttg gat gat tgg aaa gat gtc      384
Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val
      115           120           125

```

ctc cca caa ggt ata agt tgc ttt att gat gat att cag gct ggt aac 432  
 Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn  
 130 135 140

tac tcc aca tct tct tta gat gat ctc agt gaa aat ttt gcc gtt ggt 480  
 Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly  
 145 150 155 160

aaa caa ctc tta cgt gat tat aat atc gag gcc aaa cat cct gtt gta 528  
 Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val  
 165 170 175

atg gtt cct ggt gtc att tct acg gga att gaa agc tgg gga gtt att 576  
 Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile  
 180 185 190

gga gac gat gag tgc gat agt tct gcg cat ttt cgt aaa cgg ctg tgg 624  
 Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp  
 195 200 205

gga agt ttt tac atg ctg aga aca atg gtt atg gat aaa gtt tgt tgg 672  
 Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp  
 210 215 220

ttg aaa cat gta atg tta gat cct gaa aca ggt ctg gac cca ccg aac 720  
 Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn  
 225 230 235 240

ttt acg cta cgt gca gca cag ggc ttc gaa tca act gat tat ttc atc 768  
 Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile  
 245 250 255

gca ggg tat tgg att tgg aac aaa gtt ttc caa aat ctg gga gta att 816  
 Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile  
 260 265 270

ggc tat gaa ccc aat aaa atg acg agt gct gcg tat gat tgg agg ctt 864  
 Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu  
 275 280 285

gca tat tta gat cta gaa aga cgc gat agg tac ttt acg aag cta aag 912  
 Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys  
 290 295 300

gaa caa atc gaa ctg ttt cat caa tgc agt ggt gaa aaa gtt tgt tta 960  
 Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu  
 305 310 315 320

att gga cat tct atg ggt tct cag acc acc ttc tac ttt atg aaa tgg 1008  
 Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp  
 325 330 335

gtc gag gct gaa ggc cct ctc tac ggt aat ggt ggt cgt gcc tgg gtt 1056  
 Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val  
 340 345 350

```

aac gaa cac ata gat tca ttc att aat gca gca ggg acg ctt ctg ggc 1104
Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly
      355                      360                      365

gct cca aag gca gtt cca gct cta att agt ggt gaa atg aaa gat acc 1152
Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr
      370                      375                      380

att caa tta aat acg tta gcc atg tat ggt ttg gaa aag ttc ttc tca 1200
Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser
      385                      390                      395                      400

aga att gag aga gta aaa atg tta caa acg tgg ggt ggt ata cca tca 1248
Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser
      405                      410                      415

atg cta cca aag gga gaa gag gtc att tgg ggg gat atg aag tca tct 1296
Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser
      420                      425                      430

tca gag gat gca ttg aat aac aac act gac aca tac ggc aat ttc att 1344
Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile
      435                      440                      445

cga ttt gaa agg aat acg agc gat gct ttc aac aaa aat ttg aca atg 1392
Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met
      450                      455                      460

aaa gac gcc att aac atg aca tta tgg ata tca cct gaa tgg ctc caa 1440
Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln
      465                      470                      475                      480

aga aga gta cat gag cag tac tgg ttc ggc tat tcc aag aat gaa gaa 1488
Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu
      485                      490                      495

gag tta aga aaa aat gag cta cac cac aag cac tgg tgg aat cca atg 1536
Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met
      500                      505                      510

gaa gta cca ctt cca gaa gct ccc cac atg aaa atc tat tgt ata tac 1584
Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr
      515                      520                      525

ggg gtg aac aac cca act gaa agg gca tat gta tat aag gaa gag gat 1632
Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp
      530                      535                      540

gac tcc tct gct ctg aat ttg acc atc gac tac gaa agc aag caa cct 1680
Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro
      545                      550                      555                      560

gta ttc ctc acc gag ggg gac gga acc gtt ccg ctc gtg gcg cat tca 1728
Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser
      565                      570                      575

atg tgt cac aaa tgg gcc cag ggt gct tca ccg tac aac cct gcc gga 1776
Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly
      580                      585                      590

```

[illegible]



<210> 2  
<211> 661  
<212> PRT  
<213> *Saccharomyces cerevisiae*

<400> 2  
Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser  
1 5 10 15  
Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg  
20 25 30  
Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly  
35 40 45  
Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg  
50 55 60  
Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu  
65 70 75 80  
Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe  
85 90 95  
Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe  
100 105 110  
Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val  
115 120 125  
Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn  
130 135 140  
Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly  
145 150 155 160  
Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val  
165 170 175  
Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile  
180 185 190  
Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp  
195 200 205  
Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp  
210 215 220  
Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn  
225 230 235 240  
Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile  
245 250 255  
Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile  
260 265 270  
Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu  
275 280 285

Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys  
290 295 300

Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu  
305 310 315 320

Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp  
325 330 335

Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val  
340 345 350

Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly  
355 360 365

Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr  
370 375 380

Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser  
385 390 395 400

Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser  
405 410 415

Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser  
420 425 430

Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile  
435 440 445

Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met  
450 455 460

Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln  
465 470 475 480

Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu  
485 490 495

Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met  
500 505 510

Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr  
515 520 525

Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp  
530 535 540

Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro  
545 550 555 560

Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser  
565 570 575

Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly  
580 585 590

Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp  
595 600 605

Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser  
610 615 620

Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp  
625 630 635 640

Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln  
645 650 655

Met Pro Phe Pro Met  
660

<210> 3  
<211> 2312  
<212> genomic DNA  
<213> Schizosaccharomyces pombe

<400> 3  
atggcgctctt ccaagaagag caaaactcat aagaaaaaga aagaagtcaa atctctctatc 60  
gactttaccaa attcaaagaa accaactcgc gctttgagtg agcaaccttc agcgctccgaa 120  
acacaatctg tttcaaataa atcaagaaaa tctaaatttg gaaaaagatt gaattttata 180  
ttgggcgcta ttttgggaat atgctggtgct ttttttttcg ctggtggaga cgacaatgct 240  
gttttcgacc ctgctacggt agataaatct gggaatatgc taggctcttc agacttggtt 300  
gatgacatta aaggatattt atcttataat gtgtttaagg atgcaccttt tactacggac 360  
aagccttcgc agtctcttag cggaaatgaa gtccaagtgt gtcttgatat gtacaatgag 420  
ggatatcgaa gtgaccatcc tgttattatg gtccctggtg ttatcagctc aggattagaa 480  
agttggctgt ttaataattg ctcgattcct tacttttaga aacgtctttg gggtagctgg 540  
tctatgctga aggcaatgtt ccttgacaag caatgctggc ttgaacattt aatgcttgat 600  
aaaaaaaccg gcttggatcc gaagggaatt aagctgcgag cagctcaggg gtttgaagca 660  
gctgattttt ttatcacggg ctattggatt tggagttaag taattgaaaa ccttgctgca 720  
attggttatg agcctaataa catgttaagt gcttcttacc attggcgggt atcatatgca 780  
aatttagagg aacgtgataa atatttttca aagttaaaaa tgttcttga gtacagcaac 840  
attgtacata agaaaaaggc agtggttgatt tctcactcca tgggttcaca ggttaactga 900  
tattttttta agtgggttga agctgagggc tacggaaatg gtggaccgac tgggttaat 960  
gatcatattg aagcatttat aaatgtgagc ctcgatggtt gtctgactac gtttctaact 1020  
tttgaataga tatcgggatc ttgatattga gcacccaaaa cagtggcagc gcttcttatc 1080  
ggtgaaatga aagatacagg tatctgaatt acattaaaca tgttaatat taattcttgc 1140  
taaccgtttt aagctcaatt gaatcagttt tcgggtctatg ggtaagcaat aaattggttga 1200  
gattctgtac taatttactg tttagtttgg aaaaattttt tcccggtctt gaggatatt 1260  
caaaaatata aatgtgtctt actttttcta acttttaata gagagccatg atggttcgca 1320  
ctatgggagg agtttagttt atgcttecta aaggaggcga tgttgtatgg ggaaatgcca 1380  
gttgggttaag aaatagtggt tgttaatttt ttattaatat ttaggctcca gatgactta 1440  
atcaaacaaa tttttccaat ggtgcaatta ttcatatag agaagacatt gataaggacc 1500  
acgatgaatt tgacatagat gatgcattac aattttttaa aaatgttaca gatgacgatt 1560  
ttaaagtcat gctagcgaaa aatttatccc acggctcttg tggactgaa aaagaagtgt 1620  
taaaaaataa cgaatgccg tctaaatgga taaatccgct agaagtaaga acattaaagt 1680  
tactaaatta tactaaccca aatagactag tcttcttat gctctgata tgaaaattta 1740  
ttcggttcac ggggtcggaa aaccaactga gagaggttat tattatacta ataactctga 1800  
ggggcaacct gtcattgatt cctcggttaa tgaatggaaca aaagtggaaa atgtgagaga 1860  
atttatgttt caaacattct attaaactgt ttattagggt attgttatgg atgatggtga 1920  
cggaacctta ccaatattag ccttgggttt ggtgtgcaat aaagtctggc aaacaaaaag 1980  
gtttaarctt gctaatacaa gtatcacaaa tcatgaaatc aagcatgaac ctgctcggtt 2040  
tgatctgaga ggaggacctc gctcggcaga acacgtcgat atacttggaac attcagagct 2100  
aaatgtatgt tcattttacc ttacaaattt ctattactaa ctcttgaaat aaggaaatta 2160  
ttttaaaagt ttcacacagg catggtgact cggtaacaaa ccgttatata tcagatatcc 2220  
agtacggaca taagtcttgt agattgcaat taactaacta accgaacagg gaaataataa 2280  
atgagataaa tctcgataaa cctagaaatt aa 2312

<210> 4  
<211> 3685  
<212> genomic DNA  
<213> Arabidopsis thaliana

<400> 4  
atgcccctta ttcatecgga aaagccgacg gagaaaccat cgacgccgcc atctgaagag 60  
gtgggtgcacg atgaggattc gcaaaagaaa ccacacgaat ctcccaaatc ccaccataag 120  
aaatcgaaacg gaggagggaa gtgggtcggtc atcgattctc gttgttggtt cattgggtgt 180  
gtgtgtgttaa cctgggtggtt tcttctcttc ctttacaacg caatgcctgc gagcttccct 240  
cagtatgtaa cggagcgcaat caggggtcct ttgectgacc cggccgggtg taagctcaaa 300  
aaagaagggtc ttaaggcgaa acatcctggt gtcttcattc ctgggattgt caccggtggg 360  
ctcgagctct gggaaggcaa acaatgcgct gatgggttat ttgaaaaacg tttgtgggtt 420  
ggaacttttg gtgaagtcta caaaagggtg gctcaacaat tctcactctt cctttatatt 480  
gggattttgga ttggatctga tgagatcacg cacttggtgc ttcttcaaca tcaactcaac 540  
tttaattcca tgtttgtctg tcttactctt tacttttttt tttttttgat gtgaaacgt 600  
attttcttaa gagactattt ctgtatgtgt aaggtaagcg ttccaaggac gtaattggct 660  
tggactattt cgttttgatt gttaacttta ggatataaaa tagctgcctt ggaatttcaa 720  
gtcatcttat tgccaaatct gttgctagac atgcccata gtcggttcat aacaagttac 780  
ttcttttact gtcgtttgct gttagatttag ctttgtgtag cgtataatga agtagtgttt 840  
tatgtttgtt tgggaataga gaagttctaa ctacatctgt ggaaagtgtg ttccaggctgt 900  
gatagaggac tgttgcctta ttattcaact atgtatatgt gtaattaaag ctagtctctt 960  
tttgatcttt cagctcaatg tgccttcttc aatttttttc tcaatttcaa agtttccat 1020  
cgagtttatt cacatgtctt gaatttctgc catctctgt ctgttatcca gctttgaact 1080  
ctccccgacc ctgctatgga tatattaaaa aaaaagtggt ttgtgggttg catctttgtt 1140  
acgatctgca tcttctcttc tggcctcagt gttcatgttt ttgctatggt agagatgggc 1200  
aatgttactg ttgatggtaa cagtcgtata gttgatagta tcttaactaa tcaattatct 1260  
ctttgattca ggctctatg ttgggtggaa cacatgtcac ttgacaatga aactgggttg 1320  
gatccagctg gatttagagt tcgagctgta tcaggactcg tggctgctga ctactttgtt 1380  
cctggctact ttgtctgggc agtgctgatt gctaaccttg cacatattcg atatgaagag 1440  
aaaaaatagt acatggctgc atatgactgg cggctctctg ttccagaacac agaggttctt 1500  
ttctcatctg tcttctactt attctgttcc atgttaccgt tcttcttcca ttacttaagg 1560  
cttaaatatg ttctcargtg aattaatagg tacgtgatca gactcttagc cgtatgaaaa 1620  
gtaatataga gttgatggtt tctaccaacg gtggaaaaaa agcagttata gttccgcat 1680  
ccatgggggt cttgtatttt ctacatttta tgaagtgggt tgaggcacca gctcctcttg 1740  
gtggcggggg tgggcccagat ttggtgtgcaa agtatattaa ggcggtgatg aacattgggt 1800  
gaccattctt tgggttccca aaagctggtg cagggtcttt ctctgctgaa gcaaaggatg 1860  
ttgcagtttc caggtattga atatctgctt atacttttga tgatcagaac ctfggctctg 1920  
gaactcaaa gttattctact aaatatcaat tctaataaca ttgctatatt atcgctgcaa 1980  
ctgacattgg ttgattatct ttgctgctta tgcaactgaa actctcttga gattagacaa 2040  
atgatgaatt gataattctt acgcattgct ctgtgatgac cagtttctta gcttcgacga 2100  
taacatttgt catactgtct tttggagggc attgaatttc gctatggaaa gcgctggagc 2160  
ttccatgctt gcattcttta ccaatttagc ttattctgct tctttcaatt ttcttgata 2220  
tgcactctat gtccttttat tcttcttaat taaagactcg ttggattagt tgcctatat 2280  
gtcacttggt tcttaatat agaacctttac tcttcttgaa aattgcagag cagattgccc 2340  
aggattctta gacacgata tatttagact tcagaccttg cagcatgtaa tgagaatgac 2400  
acgcacatgg gactcaacaa tgtctatgct accgaaggga ggtgacacga tatggggcgg 2460  
gcttgatttg tcacccggaga aaggccacac ctggttggtgg aaaaagcaaa agaacaacga 2520  
aactcttggt gaagcaggtg aaacccgagt ttccaagaaa agtctgttta actatggaa 2580  
gatgatatct ttggggaaag agtagcaga ggctgcgcca tctgagatta ataatttga 2640  
ctctcgagta aggacatata aatcataata aaccttgtae atttctgat tgtatgatga 2700  
atatctgtac attttatctg gtgaagggtg ctgtcaaaag tcagagtatc ccaaatcaca 2760  
ctgtgctgta cgtgtggaca gagtaccatg acatgggaat tgcctgggac aaagctatcg 2820  
ctgagtataa ggtctacact gctggtgaag ctatagatct actacattat gttgctctta 2880  
agatgatggc cgttggtgct gctcattctt cctatggaa ttggtatgac ttggtatgac 2940  
ccaagtatca agatcccaaa tactggctca atccgttaga gacaaagtaa gtgattctt 3000  
gattccaact gtatctctcg tcttgatgca ttatcagctt ttctgttttc ggtcctgttg 3060  
gatatggtt tcagctcaaa gcttacaagg ctgattatga gcttcttcca aaaaaggctg 3120  
ctcagtaata ttgagtgct aaagttcgata catgtgactc ttgttatata accctccgtt 3180

tggtttgttc tgcctttttca gattaccgaa tgctcctgag atggaaatct actcattata 3240  
cggagtgggg ataccaacgg aacgagcata cgtatacaag cctaaccagt ctcccacag 3300  
tcgcacccc ttccagatat tcactttctgc tcacgaggag gacgaagata gctgtctgaa 3360  
agcaggagtt tacaatgtgg atggggatga aacagtaccc gtectaagtg ccgggtacat 3420  
gtgtgcacaaa gcgtggcgtg gcaagacaag attcaaccct tccggaatca agacttatat 3480  
aagagaatac aatcaccttc cgcgggctaa cctgttggaa gggcgcgga cgcagagtgg 3540  
tgcccatgtt gatatacatgg gaaactttgc ttgatcgaa galatcatga gggttgcgc 3600  
cggaggtaac gggctctgata taggacatga ccagggtccac tctggcatat ttgaatggtc 3660  
ggagcgtatt gacctgaagc tgtga 3685

<210> 5  
 <211> 2427  
 <212> cDNA  
 <213> Arabidopsis thaliana

<400> 5  
 agaaacagct ctttctctct ctgcactgat ctaacaatcc ctaactctgtg ttctaaattc 60  
 ctggacgaga tttgacaaaag tccgtatagc ttaacctggg ttaattttcaa gtgacagata 120  
 tgcccccttat tcatcggaag aagccgacgg agaaaccatc gacgcgcgca tctgaagagg 180  
 tgggtgcacga tgaggattcg caaaagaaac cacacgaatc ttccaaatcc caccataaga 240  
 aatcgaaacgg aggagggaag tggctcgtgca tcgattcttg ttgttgggtc attgggtgtg 300  
 tgtgtgtaac cgggtgggtt cttctcttcc tttacaacgc aatgcctgcg agcttccctc 360  
 agtatgtaac ggagcgaatc acgggtccct tgcctgaacc gcccggtgtt aagctcaaaa 420  
 aaagaagggtc ttaaggcgaa acatccctgt gtcttcattc ctgggattgt caccgggtggg 480  
 ctgcagcttt gggaaggcaa acaatgcgct gatgggttat ttagaaaaag tttgtggggg 540  
 ggaacttttg tggaagtcta caaaaggcct ctatgttggg tggaaacacat gtcacttgac 600  
 aatgaaactg ggttggatcc agctggatct agagttcgag ctgtatcagg actcgtggct 660  
 gctgactact ttgctcctgg ctactttgtc tgggcagtgc tgattgctaa ccttgacatc 720  
 attggatatg aagagaaaaa tatgtacatg gctgcataat actggcggct ttctgttctg 780  
 aacacagagg tacgtgatca gactcttagc cgtatgaaaa gtaatataga gttgatgggt 840  
 tctaccaacg gtggaaaaaa agcagttata gtcccgcatc ccattggggg cttgtattct 900  
 ctacatttta tgaagtgggt tgaggcacca gctcctctgg gtggcggggg tgggccagat 960  
 tgggtgtgcaa agtatattaa ggcggtgatg aacatttggt gaccattctc tgggtttcca 1020  
 aaagctgttg cagggctttt ctctgctgaa gcaaggatg ttgcagttgc cagagcgatt 1080  
 gccccaggat tcttagacac cgatatattt agacttcaga ccttgacgca tggtaatgaga 1140  
 atgacacgca catgggacct aacaatgtct atgttaccga agggagggtg cagcatatgg 1200  
 ggccgggcttg attggtcacc ggagaaaggg cacacctgtt gtgggaaaaa gcaaaagaac 1260  
 aacgaaactt gtgggtgaagc aggtgaaaaa ggagtttcca agaaaagtcc tgttaactat 1320  
 ggaaggatga tatcttttgg gaaagaagta gcagaggctg cggcatctga gatgaataat 1380  
 attgattttc gaggtgctgt caaaggtcag agtatcccaa atcacacctg tctgtacgtg 1440  
 tggacagagt accatgacat gggaattgct gggatcaaaag ctatcgctga gtataaggctc 1500  
 tacactgctg gtgaagctat agatctacta cattatgttg ctctaagat gatggcgcgt 1560  
 ggtgcccgtc atttctctta tggaaatgct gatgatttgg atgacacca gtatcaagat 1620  
 cccaaatact ggtcaaatcc gttagagaca aaattaccga atgctcctga gatggaaatc 1680  
 tactcattat acggagtggg gataccaacg gaacgagcat acgtatacaa gcttaaccag 1740  
 tctcccgaca gttgcatccc ctttcagata ttcacttctg ctcacgagga ggaagaaat 1800  
 agctgtctga aagcaggagt ttacaatgtg gatggggatg aaacagtacc cgtcttaagt 1860  
 gccgggtaca tgtgtgcaaa agcgtggcgt ggcaagacaa gattcaaccc ttcgggaatc 1920  
 aagacttata taagagaata caatcactct ccgcgggcta acctgttggg agggcgcggg 1980  
 acgcagagtg gtgcccattg tgatatcatg ggaaaccttg ctttgatcga agatatcatg 2040  
 agggttgcgg ccggaggtaa cgggtctgat ataggacatg accaggtcca ctctggcata 2100  
 tttgaatggg cggagcgtat tgacctgaag ctgtgaatat catgatctct ttaagctgtc 2160  
 ctgtcagctt atgtgaatcc aatactttga aagagagatc atcatcaat catcatcatc 2220  
 gtcatcatca tgatgtctaa ctacaaaaga agcctgagaa tgatactttg gtgcgaaatc 2280  
 ctcaataacct ctttaatatc cttattgaat gtaaatata caatcctatc taatgtttga 2340  
 acgataaacac aaaacttgct gngccatgt ttgtttgtct tggcaaaagc atcaatttgt 2400  
 gggttaaaaa aaaaaaaa aaaaaaa 2427

Met	Pro	Leu	Ile	His	Arg	Lys	Lys	Pro	Thr	Glu	Lys	Pro	Ser	Thr	Pro
1				5					10					15	
Pro	Ser	Glu	Glu	Val	Val	His	Asp	Glu	Asp	Ser	Gln	Lys	Lys	Pro	His
			20					25					30		
Glu	Ser	Ser	Lys	Ser	His	His	Lys	Lys	Ser	Asn	Gly	Gly	Gly	Lys	Trp
		35					40					45			
Ser	Cys	Ile	Asp	Ser	Cys	Cys	Trp	Phe	Ile	Gly	Cys	Val	Cys	Val	Thr
	50					55					60				
Trp	Trp	Phe	Leu	Leu	Phe	Leu	Tyr	Asn	Ala	Met	Pro	Ala	Ser	Phe	Pro
65					70					75					80
Gln	Tyr	Val	Thr	Glu	Arg	Ile	Thr	Gly	Pro	Leu	Pro	Asp	Pro	Pro	Gly
				85					90					95	
Val	Lys	Leu	Lys	Lys	Glu	Gly	Leu	Lys	Ala	Lys	His	Pro	Val	Val	Phe
			100					105					110		
Ile	Pro	Gly	Ile	Val	Thr	Gly	Cly	Leu	Glu	Leu	Trp	Glu	Gly	Lys	Gln
		115					120					125			
Cys	Ala	Asp	Gly	Leu	Phe	Arg	Lys	Arg	Leu	Trp	Gly	Gly	Thr	Phe	Gly
	130					135					140				
Glu	Val	Tyr	Lys	Arg	Pro	Leu	Cys	Trp	Val	Glu	His	Met	Ser	Leu	Asp
145					150					155					160
Asn	Glu	Thr	Gly	Leu	Asp	Pro	Ala	Gly	Ile	Arg	Val	Arg	Ala	Val	Ser
				165					170					175	
Gly	Leu	Val	Ala	Ala	Asp	Tyr	Phe	Ala	Pro	Gly	Tyr	Phe	Val	Trp	Ala
			180					185					190		
Val	Leu	Ile	Ala	Asn	Leu	Ala	His	Ile	Gly	Tyr	Glu	Glu	Lys	Asn	Met
		195					200				205				
Tyr	Met	Ala	Ala	Tyr	Asp	Trp	Arg	Leu	Ser	Phe	Gln	Asn	Thr	Glu	Val
	210					215					220				
Arg	Asp	Gln	Thr	Leu	Ser	Arg	Met	Lys	Ser	Asn	Ile	Glu	Leu	Met	Val
225					230					235					240
Ser	Thr	Asn	Gly	Gly	Lys	Lys	Ala	Val	Ile	Val	Pro	His	Ser	Met	Gly
			245					250						255	
Val	Leu	Tyr	Phe	Leu	His	Phe	Met	Lys	Trp	Val	Glu	Ala	Pro	Ala	Pro
			260					265					270		
Leu	Gly	Gly	Gly	Gly	Gly	Pro	Asp	Trp	Cys	Ala	Lys	Tyr	Ile	Lys	Ala
		275					280					285			



Val Met Asn Ile Gly Gly Pro Phe Leu Gly Val Pro Lys Ala Val Ala  
290 295 300

Gly Leu Phe Ser Ala Glu Ala Lys Asp Val Ala Val Ala Arg Ala Ile  
305 310 315 320

Ala Pro Gly Phe Leu Asp Thr Asp Ile Phe Arg Leu Gln Thr Leu Gln  
325 330 335

His Val Met Arg Met Thr Arg Thr Trp Asp Ser Thr Met Ser Met Leu  
340 345 350

Pro Lys Gly Gly Asp Thr Ile Trp Gly Gly Leu Asp Trp Ser Pro Glu  
355 360 365

Lys Gly His Thr Cys Cys Gly Lys Lys Gln Lys Asn Asn Glu Thr Cys  
370 375 380

Gly Glu Ala Gly Glu Asn Gly Val Ser Lys Lys Ser Pro Val Asn Tyr  
385 390 395 400

Gly Arg Met Ile Ser Phe Gly Lys Glu Val Ala Glu Ala Ala Pro Ser  
405 410 415

Glu Ile Asn Asn Ile Asp Phe Arg Gly Ala Val Lys Gly Gln Ser Ile  
420 425 430

Pro Asn His Thr Cys Arg Asp Val Trp Thr Glu Tyr His Asp Met Gly  
435 440 445

Ile Ala Gly Ile Lys Ala Ile Ala Glu Tyr Lys Val Tyr Thr Ala Gly  
450 455 460

Glu Ala Ile Asp Leu Leu His Tyr Val Ala Pro Lys Met Met Ala Arg  
465 470 475 480

Gly Ala Ala His Phe Ser Tyr Gly Ile Ala Asp Asp Leu Asp Asp Thr  
485 490 495

Lys Tyr Gln Asp Pro Lys Tyr Trp Ser Asn Pro Leu Glu Thr Lys Leu  
500 505 510

Pro Asn Ala Pro Glu Met Glu Ile Tyr Ser Leu Tyr Gly Val Gly Ile  
515 520 525

Pro Thr Glu Arg Ala Tyr Val Tyr Lys Leu Asn Gln Ser Pro Asp Ser  
530 535 540

Cys Ile Pro Phe Gln Ile Phe Thr Ser Ala His Glu Glu Asp Glu Asp  
545 550 555 560

Ser Cys Leu Lys Ala Gly Val Tyr Asn Val Asp Gly Asp Glu Thr Val  
565 570 575

Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Ala Trp Arg Gly Lys  
580 585 590

Thr Arg Phe Asn Pro Ser Gly Ile Lys Thr Tyr Ile Arg Glu Tyr Asn  
595 600 605

His Ser Pro Pro Ala Asn Leu Leu Glu Gly Arg Gly Thr Gln Ser Gly  
610 615 620

Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile Glu Asp Ile Met  
625 630 635 640

Arg Val Ala Ala Gly Gly Asn Gly Ser Asp Ile Gly His Asp Gln Val  
645 650 655

His Ser Gly Ile Phe Glu Trp Ser Glu Arg Ile Asp Leu Lys Leu  
660 665 670

<210> 7  
<211> 643  
<212> cDNA  
<213> Zea mays  
  
<220>  
<221> CDS  
<222> (1)..(402)

<400> 7  
cgg gag aaa ata gct gct ttg aag ggg ggt gtt tac tta gcc gat ggt 48  
Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly  
1 5 10 15  
gat gaa act gtt cca gtt ctt agt gcg ggc tac atg tgt gcg aaa gga 96  
Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly  
20 25 30  
tgg cgt ggc aaa act cgt ttc agc cct gcc ggc agc aag act tac gtg 144  
Trp Arg Gly Lys Thr Arg Phe Ser Pro Ala Gly Ser Lys Thr Tyr Val  
35 40 45  
aga gaa tac agc cat tcg cca ccc tct act ctc ctg gaa ggc agg ggc 192  
Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly  
50 55 60  
acc cag agc ggt gca cat gtt gat ata atg ggg aac ttt gct cta att 240  
Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile  
65 70 75 80  
gag gac gtc atc aga ata gct gct ggg gca acc ggt gag gaa att ggt 288  
Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly  
85 90 95  
ggc gat cag gtt tat tca gat ata ttc aag tgg tca gag aaa atc aaa 336  
Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Lys  
100 105 110  
ttg aaa ttg taa cct atg gga agt taa aga agt gcc gac ccg ttt att 384  
Leu Lys Leu  
115  
gcg ttc caa agt gtc ctg cctgagtgca actctggatt ttgcttaaatt 432  
attgtaattt ttcacgcttc attcgteccct ttgtcaaatt tacatttgac aggacgccaa 492  
tgcgatacga tgttgtaccg ctattttcag catgtatat taaactgtac aggtgtaagt 552  
tgcatttgcc agctgaaatt gtgtagtcgt tttctttacg atttaaatanc aagtggcgga 612  
gcagtgcgcc aagcnaaaaa aaaaaaaaaa a 643

<210> 8  
<211> 115  
<212> PRT  
<213> Zea mays

<400> 8  
Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly  
1 5 10 15  
Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly  
20 25 30  
Trp Arg Gly Lys Thr Arg Phe Ser Pro Ala Gly Ser Lys Thr Tyr Val  
35 40 45  
Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly  
50 55 60  
Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile  
65 70 75 80  
Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly  
85 90 95  
Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Lys  
100 105 110  
Leu Lys Leu  
115

<210> 9  
<211> 616  
<212> cDNA  
<213> *Neurospora crassa*

<400> 9  
ggcggcgaag acgaggcgg aagttggagg ctaacgagaa tgacnctcgg agatggatct 60  
accctctaga gacacgacta ccnttgcacc cagcctcaag gtntacngtt tntatgggta 120  
ggaagccgac ggagcgagcc tacatctatc tggcgcccga tcccgggacg acaacgcac 180  
tttagatgac gatcgatacg actttgactn aggggcacat tgaccacggt gtgattttgg 240  
gcgaaggcga tggcacagtg aaccttatga gtttggggta cctgtgcaat aaggggtgga 300  
aaatgaagag atacaatcct gcgggctcaa aaataaccgt ggtcgagatg ccgcatgaac 360  
cagaacggtt caatccgaga ggagggccga atacggcgga tcacgtggat attctaggaa 420  
ggcagaatct aaacgagtac attcttaaag tggcggcagg tcgaggcgat acaattgagg 480  
attttattac tagtaatat cttaaatatg tagaaaaggt tgaaatttat gaagagtaat 540  
taaatccggc acataggtta ctcaatagta cgactaatta aaaaaaatt ttttttctaa 600  
aaaaaaaaa aaaaaa 616

<210> 10  
<211> 1562  
<212> genomic DNA  
<213> Arabidopsis thaliana

<400> 10  
atgaaaaaaa tatcttcaca ttatttcggtg gtcatagcga tactcgttgt ggtgacgatg 60  
acctcgatgt gtcaagctgt gggtagcaac gtgtaccctt tgattcttgt tccaggaaac 120  
ggaggtaacc agctagaggt acggctggac agagaatata agccaagtag tgtctgttgt 180  
agcagctggt tatatccgat tcataagaag agtgggtgat ggtttaggct atggttcgat 240  
gcagcagtgt tattgtctcc ctccaccagg tgcttcagcg atcgaatgat gttgtactat 300  
gaccttgatt tggatgatta ccaaaatgct cctgggtgtcc aaaccgggt tccctatttc 360  
ggttcgacca aatcacttct atacctcgac cctcgtctcc ggtagtact tccaagata 420  
tatcattttg ggacatttgc ataatagaaca aaatagacat aaatttgggg gattatttgt 480  
atatcaatat ccatttatat gctagtcggt aatgtgagtg ttatgttagt atagttaatg 540  
tgagtgttat gtgattttcc attttaaatag aagctagaaa gttgtcgttt aataatgttg 600  
ctatgtcatg agaattataa ggacactatg taaatgtagc ttaataataa ggtttgatit 660  
gcagagatgc cacatcttac atggaacatt tggtgaaagc tctagagaaa aaatgcgggt 720  
atgttaacga ccaaacccatc ctaggagctc catatgattt caggtacggc ctggctgctt 780  
cgggccaccc gtcccggtga gctcacagt tctacaaga cctcaaacaa ttgggtggaa 840  
aaactagcag cgagaacgaa ggaaagccag tgatactctt ctcccatagc ctaggaggac 900  
ttttcgtcct ccatttctct aaccgtacca ccccttcctg gcgcgcgaag tacatcaaac 960  
actttgttgc actcgtctgc ccattgggtg ggaagatctc tcagatgaag acatttgcct 1020  
ctggcaacac actcgggtgtc cctttagtta accctttgct ggtcagacgg catcagagga 1080  
cctccgagag taaccaatgg ctacttccat ctaccaaaagt gtttcacgac agaactaac 1140  
cgctcgtcgt aactccccag gtttaactaca cagcttacga gatggatcgg ttttttcgag 1200  
acatcggatt ctcaacaagga gttgtgcctt acsagacaag agtgttgctt ttaacagagg 1260  
agctgatgac tccgggagtg ccagtcactt gcataatatg gagaggagt gatacacagg 1320  
aggttttgat gtatggaaaa ggaggattcg ataagcaacc agagattaaag tatggagatg 1380  
gagatgggac ggttaatttg gcgagcttag cacccttgaa agtcgatagc ttgaacacgg 1440  
tagagattga tggagtctcg catacatcta tacttaaaga cgagatcgca cttaaagaga 1500  
ttatgaagca gatttcaatt attaatatg aattagccaa tgttaattgcc gtcaatgaat 1562  
ga

<210> 11  
<211> 3896  
<212> genomic DNA  
<213> *Arabidopsis thaliana*

<400> 11  
atgggagcga attcgaatc agtaacggct tcttccaccg tcatcgccgt tttttttcttg 60  
atttgccgtg gccgaactgc ggtggaggat gagaccgagt ttcacggcga ctactcgaag 120  
ctatcgggta taatcattcc gggatttgcg tcgacgcagc tacgagccgtg gcgatccctt 180  
gaactgtccat acactccgtt ggacttcaat ccgctcgacc tcgtatggct agacaccact 240  
aagggtccgtg atcttcattt ccttcgctcc ttattctgtc ggtcgagtca cttgttcatg 300  
aatcccaagc gaaatatagc aatgaagcat gtctcgtctc tcttatcgat tcgttcatta 360  
gtcaacagtg acgcttctga atctgagttt agagtcatat aaaacagctg actcggcgag 420  
tgtttcccat cgttttttgt tcgctaaatg tagcgcaatg aatgtgtaat tagtctgcgc 480  
tttttattca actagatctg caagtttttc agagtgtcga atagtagtta gaaaatgta 540  
gggtcatttta cttgtgcatt gtgattcttt ttgttggtgc ttactgatcg acgtgatgga 600  
tggttttacag cttctttctg ctgccaactg ctggtttaag tgtatgggtg tagatcccta 660  
taatcaaaaca gaccatcccg agtgaagtc accgctcgac agtggctctt cagccatcac 720  
agaatcggat ccagggttaca taacaggtag ttccggattt ttctttctct tgagttttct 780  
tcaatttgat atcatcttgt tgtgatataa tatggctaag ttcattaatt tgggtcaatt 840  
tcaggctctc tttctactgt ctggaagag tggtctaagt ggtgtgttga gtttgggtata 900  
gaagcaaatg caattgtcgc tgttccatgc gattggagat tgtccccaac caaattggaa 960  
gagcgtgacc tttactttca caagctcaag ttagtcttta tcaggctaat gtctttctatc 1020  
ttctcttttt atgtaagata agctaagagc tctggtctgt ttcttttttg caggttgacc 1080  
tttgaactctg ctttaaaact ccgtggcgcc ccttctatag tatttgcca ttcattgggt 1140  
aataatgtct tcagatactt tctggaatgg ctgaggctag aaattgcacc aaaacattat 1200  
ttgaagtggc ttgatcagca tatccatgct tacttcgctg ttgggtaccg cctactatcc 1260  
ttaagttacc attttatctt tctctcaatt gggggagtta tgtgtgact tactggattg 1320  
agctcgatac ctgatttgtt gtgatcttag gagctccctt tcttgggtct gttgaggcaa 1380  
tcaaatctac tctctctggg gtaacgtttg gccttctctt tcttgaggtg acctctgact 1440  
tctctttagt ttaagtagt tgatatcaac caggtcttat aactcactgg attttccttt 1500  
tgaaagtatt acttctgtta attgaactgc tgtacgcgat atgggtatct tagatcttga 1560  
agtgttagtt atcaaagaac atatttgtgg tagtatacct gtcagcggcc tttagctaata 1620  
caaccaaac acatgtacac tgatttagtt ttcagattat tatggttagac tttaagttga 1680  
gaagaaaact tgactgaact ctttttattt taataggcta tgatttggct attgaaatca 1740  
tgtgacatat tgacatgcgc ttctcatgtt ttttgttggc aaggcttcag ggaactgctc 1800  
ggtgtttgtc caattctttt cgtctgtcat tctggcttat gccatttca aagaattgca 1860  
aggttgataa cacattctgg acgcattttt ctgggggtgc tgcaaaagaa gatzagcgcg 1920  
tataccactg tgatgaagag gaatatcaat caaaatattc tggctggccg acaaatatta 1980  
ttaacattga aattctctcc actagcggtt agactctgta tatgcaactg taacactaac 2040  
aaaagtcca ccaagaatgt tcactcccat atttcgttcc ttgatgtgt atccatcagt 2100  
tacagaaaac gctctagtca acatgaccag catggaatgt ggcttccca ccttttctc 2160  
tttcacagcc cgtgaactag cagatgggac ttttttcaaa gcaatagaag acctatgacc 2220  
agatagcaag aggatgttac accagctaaa gaagtacgta cctttctttg tgataagaaa 2280  
tattgtctat cgatcatcac ttgctggctt cttgtacgtc aaattgtttt gtttaaatct 2340  
ctatatcaat tgttcatatg ctttgtcttt cttaactataa gaaacaagta taatcagaaa 2400  
ccttattatt gattatcagt tctctcttta tattatggaa tgcctttctc gtttacagtt 2460  
atgaatgcaa aagggggtat tttagtcgat tgattctctc attctctagt ttgttttgac 2520  
taatagcgtc aattttgttt ttctagcaaa tctttgtgaa ttatatataa catgctaact 2580  
atacttttca ggttgatca tgatgacctt gtttttaac ccttgactcc ttgggagaga 2640  
ccacctataa aaaatgtatt ttgcatacat ggtgtctcat taagacaga ggtatgatgc 2700  
attctcaata tcacattatg cgttgacttt gttattatat ccccatcttg gtttgcaata 2760  
cttttttgaa ttatgattta cttctctcct tgcattctat gctattraag gttaaaagga 2820  
ctaaatgtat gaagcgtct gtcataggtt ggttattact ttgccccag tggcaaacct 2880  
tattctgata attggatcac caccgatata atttatgaaa ctgaaggttc cctcgtgtca 2940  
aggtaatctt ccgcaatggc agaagtaaaa caggaaggca aagctctctg tatcagttca 3000  
gtggcatgtt atctcagttg cataagcaaa ttattaaaca actaaaattt aagtactctt 3060  
ctatcatccc ttttgagctt agtggatgat cagtggctta aagtggaag aggtgttgca 3120  
tgaaacatga cacttgtatc aaagataact agcaaaacaa aactaaccca tttctgaatt 3180

tcataattact aggagtagtc gtgcttttaa aaaatttgtt ttaagaaacc gaaaaactag 3240  
ttcataatcctt gattgtgcaa tatctgcagg tctggaactg tggttgatgg gaacgctgga 3300  
cctataactg gggatgagac ggtaagctca gaagtgtggt ttgaaattat cttcttgcaa 3360  
actactgaag actaagataa tacttgcttc tggaaactg ctgctatgt tctctagtac 3420  
actgcaatat tgactctccg ctacttttat cgattatgaa attgatctct tatagggtacc 3480  
ctatcattca ctctcttggg gcaagsattg gctcggacct aaagttaaca taacaatggc 3540  
tcccaggtta ctctttttta gttectcacc ttatatagat caaactttta gtgtactttt 3600  
ctgggttatgt gttgatttac ctccaatttg ttctttctaa aaatcatata tctctgtact 3660  
cctcaagaac ttgtattaat ctaaacgaga ttctcattgg gaaaataaaa caacagccag 3720  
aacacgatgg aagcgacgta catgtggaac taaatgttga tcatgagcat gggtcagaca 3780  
tcatagctaa catgacaaaa gcaccaaggg ttaagtacat aaccttttat gaagactctg 3840  
agagcattec ggggaagaga accgcagctc gggagcttga taaaagtggg tattaa 3896



<210> 12  
<211> 709  
<212> cDNA  
<213> *Lycopersicon esculentum*

<400> 12  
ctggggccaa aagtgaacat aacaaggaca ccacagtcag agcatgatgt tcagatgtac 60  
aagtgcacat aaatatagag catcaacatg gtgaagatat cattcccaat atgacaaagt 120  
tacctacaat gaagtacata acctattatg aggattctga aagttttcca gggacaagaa 180  
cagcagtttg ggagcttgat aaagcaaate acaggaacat tgtcagatct ccagctttga 240  
tgcgggagct gtggcttgag atgtggcatg atattcatcc tgataaaaag tccaagtttg 300  
ttacaaaagg tgggtgtctga tcttcactat tttcttctat aaatgtttga gtttgtattg 360  
acattgtaag tattgcaaca aaaagcaaag cgtgggcctc tgagggatga ggactgctat 420  
tgggattacg ggaaagctcg atgtgcatgg gctgaacatt gtgaatacag gttagaatat 480  
tcaaattata ttttgcaaaa ttttctcttt ttgtgtattt aggccacctt tccccgggtca 540  
caacgatgca gatatgtatt cggggatggt caccggggac agagttgcag attgaagagt 600  
tctacatctc acatcctgtc acactatgtg tgatatttaa gaaactttgt ttggcggaac 660  
aacaagtttg cacaacatt tgaagaagaa agcgaaatga ttcagagag 709

<210> .13  
<211> 623  
<212> PRT  
<213> Schizosaccharomyces pombe

<400> 13

Met	Ala	Ser	Ser	Lys	Lys	Ser	Lys	Thr	His	Lys	Lys	Lys	Lys	Glu	Val
1				5				10						15	
Lys	Ser	Pro	Ile	Asp	Leu	Pro	Asn	Ser	Lys	Lys	Pro	Thr	Arg	Ala	Leu
			20					25					30		
Ser	Glu	Gln	Pro	Ser	Ala	Ser	Glu	Thr	Gln	Ser	Val	Ser	Asn	Lys	Ser
		35					40					45			
Arg	Lys	Ser	Lys	Phe	Gly	Lys	Arg	Leu	Asn	Phe	Ile	Leu	Gly	Ala	Ile
	50					55					60				
Leu	Gly	Ile	Cys	Gly	Ala	Phe	Phe	Phe	Ala	Val	Gly	Asp	Asp	Asn	Ala
	65				70					75					80
Val	Phe	Asp	Pro	Ala	Thr	Leu	Asp	Lys	Phe	Gly	Asn	Met	Leu	Gly	Ser
				85					90					95	
Ser	Asp	Leu	Phe	Asp	Asp	Ile	Lys	Gly	Tyr	Leu	Ser	Tyr	Asn	Val	Phe
		100						105					110		
Lys	Asp	Ala	Pro	Phe	Thr	Thr	Asp	Lys	Pro	Ser	Gln	Ser	Pro	Ser	Gly
		115					120					125			
Asn	Glu	Val	Gln	Val	Gly	Leu	Asp	Met	Tyr	Asn	Glu	Gly	Tyr	Arg	Ser
	130					135					140				
Asp	His	Pro	Val	Ile	Met	Val	Pro	Gly	Val	Ile	Ser	Ser	Gly	Leu	Glu
	145				150					155				160	
Ser	Trp	Ser	Phe	Asn	Asn	Cys	Ser	Ile	Pro	Tyr	Phe	Arg	Lys	Arg	Leu
				165					170					175	
Trp	Gly	Ser	Trp	Ser	Met	Leu	Lys	Ala	Met	Phe	Leu	Asp	Lys	Gln	Cys
		180						185					190		
Trp	Leu	Glu	His	Leu	Met	Leu	Asp	Lys	Lys	Thr	Gly	Leu	Asp	Pro	Lys
		195					200					205			
Gly	Ile	Lys	Leu	Arg	Ala	Ala	Gln	Gly	Phe	Glu	Ala	Ala	Asp	Phe	Phe
	210						215					220			
Ile	Thr	Gly	Tyr	Trp	Ile	Trp	Ser	Lys	Val	Ile	Glu	Asn	Leu	Ala	Ala
	225				230					235				240	
Ile	Gly	Tyr	Glu	Pro	Asn	Asn	Met	Leu	Ser	Ala	Ser	Tyr	Asp	Trp	Arg
				245					250					255	
Leu	Ser	Tyr	Ala	Asn	Leu	Glu	Glu	Arg	Asp	Lys	Tyr	Phe	Ser	Lys	Leu
			260					265					270		

Lys Met Phe Ile Glu Tyr Ser Asn Ile Val His Lys Lys Lys Val Val  
275 280 285

Leu Ile Ser His Ser Met Gly Ser Gln Val Thr Tyr Tyr Phe Phe Lys  
290 295 300

Trp Val Glu Ala Glu Gly Tyr Gly Asn Gly Gly Pro Thr Trp Val Asn  
305 310 315 320

Asp His Ile Glu Ala Phe Ile Asn Ile Ser Gly Ser Leu Ile Gly Ala  
325 330 335

Pro Lys Thr Val Ala Ala Leu Leu Ser Gly Glu Met Lys Asp Thr Gly  
340 345 350

Ile Val Ile Thr Leu Asn Ile Leu Glu Lys Phe Phe Ser Arg Ser Glu  
355 360 365

Arg Ala Met Met Val Arg Thr Met Gly Gly Val Ser Ser Met Leu Pro  
370 375 380

Lys Gly Gly Asp Val Ala Pro Asp Asp Leu Asn Gln Thr Asn Phe Ser  
385 390 395 400

Asn Gly Ala Ile Ile Arg Tyr Arg Glu Asp Ile Asp Lys Asp His Asp  
405 410 415

Glu Phe Asp Ile Asp Asp Ala Leu Gln Phe Leu Lys Asn Val Thr Asp  
420 425 430

Asp Asp Phe Lys Val Met Leu Ala Lys Asn Tyr Ser His Gly Leu Ala  
435 440 445

Trp Thr Glu Lys Glu Val Leu Lys Asn Asn Glu Met Pro Ser Lys Trp  
450 455 460

Ile Asn Pro Leu Glu Thr Ser Leu Pro Tyr Ala Pro Asp Met Lys Ile  
465 470 475 480

Tyr Cys Val His Gly Val Gly Lys Pro Thr Glu Arg Gly Tyr Tyr Tyr  
485 490 495

Thr Asn Asn Pro Glu Gly Gln Pro Val Ile Asp Ser Ser Val Asn Asp  
500 505 510

Gly Thr Lys Val Glu Asn Gly Ile Val Met Asp Asp Gly Asp Gly Thr  
515 520 525

Leu Pro Ile Leu Ala Leu Gly Leu Val Cys Asn Lys Val Trp Gln Thr  
530 535 540

Lys Arg Phe Asn Pro Ala Asn Thr Ser Ile Thr Asn Tyr Glu Ile Lys  
545 550 555 560

His Glu Pro Ala Ala Phe Asp Leu Arg Gly Gly Pro Arg Ser Ala Glu  
565 570 575

His Val Asp Ile Leu Gly His Ser Glu Leu Asn Glu Ile Ile Leu Lys  
580 585 590

Val Ser Ser Gly His Gly Asp Ser Val Pro Asn Arg Tyr Ile Ser Asp  
595 600 605

Ile Gln Glu Ile Ile Asn Glu Ile Asn Leu Asp Lys Pro Arg Asn  
610 615 620

<210> 14  
<211> 432  
<212> PRT  
<213> Arabidopsis thaliana

<400> 14  
Met Lys Lys Ile Ser Ser His Tyr Ser Val Val Ile Ala Ile Leu Val  
1 5 10 15  
Val Val Thr Met Thr Ser Met Cys Gln Ala Val Gly Ser Asn Val Tyr  
20 25 30  
Pro Leu Ile Leu Val Pro Gly Asn Gly Gly Asn Gln Leu Glu Val Arg  
35 40 45  
Leu Asp Arg Glu Tyr Lys Pro Ser Ser Val Trp Cys Ser Ser Trp Leu  
50 55 60  
Tyr Pro Ile His Lys Lys Ser Gly Gly Trp Phe Arg Leu Trp Phe Asp  
65 70 75 80  
Ala Ala Val Leu Leu Ser Pro Phe Thr Arg Cys Phe Ser Asp Arg Met  
85 90 95  
Met Leu Tyr Tyr Asp Pro Asp Leu Asp Asp Tyr Gln Asn Ala Pro Gly  
100 105 110  
Val Gln Thr Arg Val Pro His Phe Gly Ser Thr Lys Ser Leu Leu Tyr  
115 120 125  
Leu Asp Pro Arg Leu Arg Asp Ala Thr Ser Tyr Met Glu His Leu Val  
130 135 140  
Lys Ala Leu Glu Lys Lys Cys Gly Tyr Val Asn Asp Gln Thr Ile Leu  
145 150 155 160  
Gly Ala Pro Tyr Asp Phe Arg Tyr Gly Leu Ala Ala Ser Gly His Pro  
165 170 175  
Ser Arg Val Ala Ser Gln Phe Leu Gln Asp Leu Lys Gln Leu Val Glu  
180 185 190  
Lys Thr Ser Ser Glu Asn Glu Gly Lys Pro Val Ile Leu Leu Ser His  
195 200 205  
Ser Leu Gly Gly Leu Phe Val Leu His Phe Leu Asn Arg Thr Thr Pro  
210 215 220  
Ser Trp Arg Arg Lys Tyr Ile Lys His Phe Val Ala Leu Ala Ala Pro  
225 230 235 240  
Trp Gly Gly Thr Ile Ser Gln Met Lys Thr Phe Ala Ser Gly Asn Thr  
245 250 255  
Leu Gly Val Pro Leu Val Asn Pro Leu Leu Val Arg Arg His Gln Arg  
260 265 270  
Thr Ser Glu Ser Asn Gln Trp Leu Leu Pro Ser Thr Lys Val Phe His  
275 280 285

Asp Arg Thr Lys Pro Leu Val Val Thr Pro Gln Val Asn Tyr Thr Ala  
290 295 300

Tyr Glu Met Asp Arg Phe Phe Ala Asp Ile Gly Phe Ser Gln Gly Val  
305 310 315 320

Val Pro Tyr Lys Thr Arg Val Leu Pro Leu Thr Glu Glu Leu Met Thr  
325 330 335

Pro Gly Val Pro Val Thr Cys Ile Tyr Gly Arg Gly Val Asp Thr Pro  
340 345 350

Glu Val Leu Met Tyr Gly Lys Gly Gly Phe Asp Lys Gln Pro Glu Ile  
355 360 365

Lys Tyr Gly Asp Gly Asp Gly Thr Val Asn Leu Ala Ser Leu Ala Ala  
370 375 380

Leu Lys Val Asp Ser Leu Asn Thr Val Glu Ile Asp Gly Val Ser His  
385 390 395 400

Thr Ser Ile Leu Lys Asp Glu Ile Ala Leu Lys Glu Ile Met Lys Gln  
405 410 415

Ile Ser Ile Ile Asn Tyr Glu Leu Ala Asn Val Asn Ala Val Asn Glu  
420 425 430

<210> 15  
<211> 552  
<212> PRT  
<213> Arabidopsis thaliana

<400> 15  
Met Gly Ala Asn Ser Lys Ser Val Thr Ala Ser Phe Thr Val Ile Ala  
1 5 10 15  
Val Phe Phe Leu Ile Cys Gly Gly Arg Thr Ala Val Glu Asp Glu Thr  
20 25 30  
Glu Phe His Gly Asp Tyr Ser Lys Leu Ser Gly Ile Ile Ile Pro Gly  
35 40 45  
Phe Ala Ser Thr Gln Leu Arg Ala Trp Ser Ile Leu Asp Cys Pro Tyr  
50 55 60  
Thr Pro Leu Asp Phe Asn Pro Leu Asp Leu Val Trp Leu Asp Thr Thr  
65 70 75 80  
Lys Leu Leu Ser Ala Val Asn Cys Trp Phe Lys Cys Met Val Leu Asp  
85 90 95  
Pro Tyr Asn Gln Thr Asp His Pro Glu Cys Lys Ser Arg Pro Asp Ser  
100 105 110  
Gly Leu Ser Ala Ile Thr Glu Leu Asp Pro Gly Tyr Ile Thr Gly Pro  
115 120 125  
Leu Ser Thr Val Trp Lys Glu Trp Leu Lys Trp Cys Val Glu Phe Gly  
130 135 140  
Ile Glu Ala Asn Ala Ile Val Ala Val Pro Tyr Asp Trp Arg Leu Ser  
145 150 155 160  
Pro Thr Lys Leu Glu Glu Arg Asp Leu Tyr Phe His Lys Leu Lys Leu  
165 170 175  
Thr Phe Glu Thr Ala Leu Lys Leu Arg Gly Gly Pro Ser Ile Val Phe  
180 185 190  
Ala His Ser Met Gly Asn Asn Val Phe Arg Tyr Phe Leu Glu Trp Leu  
195 200 205  
Arg Leu Glu Ile Ala Pro Lys His Tyr Leu Lys Trp Leu Asp Gln His  
210 215 220  
Ile His Ala Tyr Phe Ala Val Gly Ala Pro Leu Leu Gly Ser Val Glu  
225 230 235 240  
Ala Ile Lys Ser Thr Leu Ser Gly Val Thr Phe Gly Leu Pro Val Ser  
245 250 255  
Glu Gly Thr Ala Arg Leu Leu Ser Asn Ser Phe Ala Ser Ser Leu Trp  
260 265 270  
Leu Met Pro Phe Ser Lys Asn Cys Lys Gly Asp Asn Thr Phe Trp Thr  
275 280 285

His Phe Ser Gly Gly Ala Ala Lys Lys Asp Lys Arg Val Tyr His Cys  
290 295 300

Asp Glu Glu Glu Tyr Gln Ser Lys Tyr Ser Gly Trp Pro Thr Asn Ile  
305 310 315 320

Ile Asn Ile Glu Ile Pro Ser Thr Ser Ala Arg Glu Leu Ala Asp Gly  
325 330 335

Thr Leu Phe Lys Ala Ile Glu Asp Tyr Asp Pro Asp Ser Lys Arg Met  
340 345 350

Leu His Gln Leu Lys Lys Tyr Val Pro Phe Phe Val Ile Arg Asn Ile  
355 360 365

Ala His Arg Ser Ser Leu Ala Gly Phe Leu Leu Tyr His Asp Asp Pro  
370 375 380

Val Phe Asn Pro Leu Thr Pro Trp Glu Arg Pro Pro Ile Lys Asn Val  
385 390 395 400

Phe Cys Ile Tyr Gly Ala His Leu Lys Thr Glu Val Gly Tyr Tyr Phe  
405 410 415

Ala Pro Ser Gly Lys Pro Tyr Pro Asp Asn Trp Ile Ile Thr Asp Ile  
420 425 430

Ile Tyr Glu Thr Glu Gly Ser Leu Val Ser Arg Ser Gly Thr Val Val  
435 440 445

Asp Gly Asn Ala Gly Pro Ile Thr Gly Asp Glu Thr Val Pro Tyr His  
450 455 460

Ser Leu Ser Trp Cys Lys Asn Trp Leu Gly Pro Lys Val Asn Ile Thr  
465 470 475 480

Met Ala Pro Gln Ile Leu Ile Gly Lys Ile Lys Gln Gln Pro Glu His  
485 490 495

Asp Gly Ser Asp Val His Val Glu Leu Asn Val Asp His Glu His Gly  
500 505 510

Ser Asp Ile Ile Ala Asn Met Thr Lys Ala Pro Arg Val Lys Tyr Ile  
515 520 525

Thr Phe Tyr Glu Asp Ser Glu Ser Ile Pro Gly Lys Arg Thr Ala Val  
530 535 540

Trp Glu Leu Asp Lys Ser Gly Tyr  
545 550



<210> 16  
 <211> 661  
 <212> PRT  
 <213> *Saccharomyces cerevisiae*

<400> 16  
 Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser  
     1                    5                    10                    15  
 Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg  
                     20                    25                    30  
 Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly  
                     35                    40                    45  
 Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg  
                     50                    55                    60  
 Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu  
                     65                    70                    75                    80  
 Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe  
                     85                    90                    95  
 Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe  
                     100                    105                    110  
 Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val  
                     115                    120                    125  
 Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn  
                     130                    135                    140  
 Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly  
                     145                    150                    155                    160  
 Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val  
                     165                    170                    175  
 Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile  
                     180                    185                    190  
 Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp  
                     195                    200                    205  
 Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp  
                     210                    215                    220  
 Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn  
                     225                    230                    235                    240  
 Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile  
                     245                    250                    255  
 Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile  
                     260                    265                    270  
 Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu  
                     275                    280                    285

Ala	Tyr	Leu	Asp	Leu	Glu	Arg	Arg	Asp	Arg	Tyr	Phe	Thr	Lys	Leu	Lys
290						295					300				
Glu	Gln	Ile	Glu	Leu	Phe	His	Gln	Leu	Ser	Gly	Glu	Lys	Val	Cys	Leu
305					310					315					320
Ile	Gly	His	Ser	Met	Gly	Ser	Gln	Ile	Ile	Phe	Tyr	Phe	Met	Lys	Trp
				325					330					335	
Val	Glu	Ala	Glu	Gly	Pro	Leu	Tyr	Gly	Asn	Gly	Gly	Arg	Gly	Trp	Val
			340					345					350		
Asn	Glu	His	Ile	Asp	Ser	Phe	Ile	Asn	Ala	Ala	Gly	Thr	Leu	Leu	Gly
	355					360					365				
Ala	Pro	Lys	Ala	Val	Pro	Ala	Leu	Ile	Ser	Gly	Glu	Met	Lys	Asp	Thr
	370					375					380				
Ile	Gln	Leu	Asn	Thr	Leu	Ala	Met	Tyr	Gly	Leu	Glu	Lys	Phe	Phe	Ser
385					390					395					400
Arg	Ile	Glu	Arg	Val	Lys	Met	Leu	Gln	Thr	Trp	Gly	Gly	Ile	Pro	Ser
			405						410					415	
Met	Leu	Pro	Lys	Gly	Glu	Glu	Val	Ile	Trp	Gly	Asp	Met	Lys	Ser	Ser
			420					425					430		
Ser	Glu	Asp	Ala	Leu	Asn	Asn	Asn	Thr	Asp	Thr	Tyr	Gly	Asn	Phe	Ile
		435				440					445				
Arg	Phe	Glu	Arg	Asn	Thr	Ser	Asp	Ala	Phe	Asn	Lys	Asn	Leu	Thr	Met
	450					455					460				
Lys	Asp	Ala	Ile	Asn	Met	Thr	Leu	Ser	Ile	Ser	Pro	Glu	Trp	Leu	Gln
465					470					475					480
Arg	Arg	Val	His	Glu	Gln	Tyr	Ser	Phe	Gly	Tyr	Ser	Lys	Asn	Glu	Glu
			485						490					495	
Glu	Leu	Arg	Lys	Asn	Glu	Leu	His	His	Lys	His	Trp	Ser	Asn	Pro	Met
			500					505					510		
Glu	Val	Pro	Leu	Pro	Glu	Ala	Pro	His	Met	Lys	Ile	Tyr	Cys	Ile	Tyr
		515					520				525				
Gly	Val	Asn	Asn	Pro	Thr	Glu	Arg	Ala	Tyr	Val	Tyr	Lys	Glu	Glu	Asp
	530					535					540				
Asp	Ser	Ser	Ala	Leu	Asn	Leu	Thr	Ile	Asp	Tyr	Glu	Ser	Lys	Gln	Pro
545					550				555						560
Val	Phe	Leu	Thr	Glu	Gly	Asp	Gly	Thr	Val	Pro	Leu	Val	Ala	His	Ser
			565						570				575		
Met	Cys	His	Lys	Trp	Ala	Gln	Gly	Ala	Ser	Pro	Tyr	Asn	Pro	Ala	Gly
			580					585					590		
Ile	Asn	Val	Thr	Ile	Val	Glu	Met	Lys	His	Gln	Pro	Asp	Arg	Phe	Asp
	595					600						605			

Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser  
610 615 620

Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp  
625 630 635 640

Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln  
645 650 655

Met Pro Phe Pro Met  
660

<210> 17  
<211> 387  
<212> PRT  
<213> Arabidopsis thaliana

<400> 17  
Val Gly Ser Asn Val Tyr Pro Leu Ile Leu Val Pro Gly Asn Gly Gly  
1 5 10 15  
Asn Gln Leu Glu Val Arg Leu Asp Arg Glu Tyr Lys Pro Ser Ser Val  
20 25 30  
Trp Cys Ser Ser Trp Leu Tyr Pro Ile His Lys Lys Ser Gly Gly Trp  
35 40 45  
Phe Arg Leu Trp Phe Asp Ala Ala Val Leu Leu Ser Pro Phe Thr Arg  
50 55 60  
Cys Phe Ser Asp Arg Met Met Leu Tyr Tyr Asp Pro Asp Leu Asp Asp  
65 70 75 80  
Tyr Gln Asn Ala Pro Gly Val Gln Thr Arg Val Pro His Phe Gly Ser  
85 90 95  
Thr Lys Ser Leu Leu Tyr Leu Asp Pro Arg Leu Arg Asp Ala Thr Ser  
100 105 110  
Tyr Met Glu His Leu Val Lys Ala Leu Glu Lys Lys Cys Gly Tyr Val  
115 120 125  
Asn Asp Gln Thr Ile Leu Gly Ala Pro Tyr Asp Phe Arg Tyr Gly Leu  
130 135 140  
Ala Ala Ser Gly His Pro Ser Arg Val Ala Ser Gln Phe Leu Gln Asp  
145 150 155 160  
Leu Lys Gln Leu Val Glu Lys Thr Ser Ser Glu Asn Glu Gly Lys Pro  
165 170 175  
Val Ile Leu Leu Ser His Ser Leu Gly Gly Leu Phe Val Leu His Phe  
180 185 190  
Leu Asn Arg Thr Thr Pro Ser Trp Arg Arg Lys Tyr Ile Lys His Phe  
195 200 205  
Val Ala Leu Ala Ala Pro Trp Gly Gly Thr Ile Ser Gln Met Lys Thr  
210 215 220  
Phe Ala Ser Gly Asn Thr Leu Gly Val Pro Leu Val Asn Pro Leu Leu  
225 230 235 240  
Val Arg Arg His Gln Arg Thr Ser Glu Ser Asn Gln Trp Leu Leu Pro  
245 250 255  
Ser Thr Lys Val Phe His Asp Arg Thr Lys Pro Leu Val Val Thr Pro  
260 265 270  
Gln Val Asn Tyr Thr Ala Tyr Glu Met Asp Arg Phe Phe Ala Asp Ile  
275 280 285

Gly Phe Ser Gln Gly Val Val Pro Tyr Lys Thr Arg Val Leu Pro Leu  
290 295 300

Thr Glu Glu Leu Met Thr Pro Gly Val Pro Val Thr Cys Ile Tyr Gly  
305 310 315 320

Arg Gly Val Asp Thr Pro Glu Val Leu Met Tyr Gly Lys Gly Gly Phe  
325 330 335

Asp Lys Gln Pro Glu Ile Lys Tyr Gly Asp Gly Asp Gly Thr Val Asn  
340 345 350

Leu Ala Ser Leu Ala Ala Leu Lys Val Asp Ser Leu Asn Thr Val Glu  
355 360 365

Ile Asp Gly Val Ser His Thr Ser Ile Leu Lys Asp Glu Ile Ala Leu  
370 375 380

Lys Glu Ile  
385

<210> 18  
<211> 389  
<212> PRT  
<213> Arabidopsis thaliana

<400> 18  
Leu Lys Lys Glu Gly Leu Lys Ala Lys His Pro Val Val Phe Ile Pro  
1 5 10 15  
Gly Ile Val Thr Gly Gly Leu Glu Leu Trp Glu Gly Lys Gln Cys Ala  
20 25 30  
Asp Gly Leu Phe Arg Lys Arg Leu Trp Gly Gly Thr Phe Leu Cys Trp  
35 40 45  
Val Glu His Met Ser Leu Asp Asn Glu Thr Gly Leu Asp Pro Ala Gly  
50 55 60  
Ile Arg Val Arg Ala Val Ser Gly Leu Val Ala Ala Asp Tyr Phe Ala  
65 70 75 80  
Pro Gly Tyr Phe Val Trp Ala Val Leu Ile Ala Asn Leu Ala His Ile  
85 90 95  
Gly Tyr Glu Glu Lys Asn Met Tyr Met Ala Ala Tyr Asp Trp Arg Leu  
100 105 110  
Ser Phe Gln Asn Thr Glu Arg Asp Gln Thr Leu Ser Arg Met Lys Ser  
115 120 125  
Asn Ile Glu Leu Met Val Ser Thr Asn Gly Gly Lys Lys Ala Val Ile  
130 135 140  
Val Pro His Ser Met Gly Val Leu Tyr Phe Leu His Phe Met Lys Trp  
145 150 155 160  
Val Glu Ala Pro Ala Pro Leu Gly Gly Gly Gly Gly Pro Asp Trp Cys  
165 170 175  
Ala Lys Tyr Ile Lys Ala Val Met Asn Ile Gly Gly Pro Phe Leu Gly  
180 185 190  
Val Pro Lys Ala Val Ala Gly Leu Phe Ser Ala Glu Ala Lys Asp Met  
195 200 205  
Arg Met Thr Arg Thr Trp Asp Ser Thr Met Ser Met Leu Pro Lys Gly  
210 215 220  
Gly Asp Thr Ile Trp Gly Gly Leu Asp Trp Ser Pro Glu Leu Pro Asn  
225 230 235 240  
Ala Pro Glu Met Glu Ile Tyr Ser Leu Tyr Gly Val Gly Ile Pro Thr  
245 250 255  
Glu Arg Ala Tyr Val Tyr Lys Leu Asn Gln Ser Pro Asp Ser Cys Ile  
260 265 270  
Pro Phe Gln Ile Phe Thr Ser Ala His Glu Glu Asp Glu Asp Ser Cys  
275 280 285

Leu Lys Ala Gly Val Tyr Asn Val Asp Gly Asp Glu Thr Val Pro Val  
290 295 300

Leu Ser Ala Gly Tyr Met Cys Ala Lys Ala Trp Arg Gly Lys Thr Arg  
305 310 315 320

Phe Asn Pro Ser Gly Ile Lys Thr Tyr Ile Arg Glu Tyr Asn His Ser  
325 330 335

Pro Pro Ala Asn Leu Leu Glu Gly Arg Gly Thr Gln Ser Gly Ala His  
340 345 350

Val Asp Ile Met Gly Asn Phe Ala Leu Ile Glu Asp Ile Met Arg Val  
355 360 365

Ala Ala Gly Gly Asn Gly Ser Asp Ile Gly His Asp Gln Val His Ser  
370 375 380

Gly Ile Phe Glu Trp  
385

<210> 19  
 <211> 1986  
 <212> DNA  
 <213> *Saccharomyces cerevisiae*

<220>  
 <221> CDS  
 <222> (1)..(1983)

<400> 19  
 atg ggc aca ctg ttt cga aga aat gtc cag aac caa aag agt gat tct 48  
 Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser  
 1 5 10 15  
 gat gaa aac aat aaa ggg ggt tct gtt cat aac aag cga gag agc aga 96  
 Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg  
 20 25 30  
 aac cac att cat cat caa cag gga tta ggc cat aag aga aga agg ggt 144  
 Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly  
 35 40 45  
 att agt ggc agt gca aaa aga aat gag cgt ggc aaa gat ttc gac agg 192  
 Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg  
 50 55 60  
 aaa aga gac ggg aac ggt aga aaa cgt tgg aga gat tcc aga aga ctg 240  
 Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu  
 65 70 75 80  
 att ttc att ctt ggt gca ttc tta ggt gta ctt ttg ccg ttt agc ttt 288  
 Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe  
 85 90 95  
 ggc gct tat cat gtc cat aat agc gat agc gac ttg ttt gac aac ttt 336  
 Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe  
 100 105 110  
 gta aat ttt gat tca ctt aaa gtg tat ttg gat gat tgg aaa gac gtt 384  
 Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val  
 115 120 125  
 ctc cca caa ggt ata agt tcg ttt att gat gat att cag gct ggt aac 432  
 Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn  
 130 135 140  
 tac tcc aca tct tct tta gat gat ctc agt gaa aat ttt gcc gtt ggt 480  
 Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly  
 145 150 155 160  
 aaa caa ctc tta cct gat tat aat atc gag gcc aaa cat cct gtt gta 528  
 Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val  
 165 170 175  
 atg gtt cct ggt gtc att tct acg gga att gaa agc tgg gga gtt att 576  
 Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile  
 180 185 190



```

gga gac gat gag tgc gat agt tct gcg cat ttt cgt aaa cgg ctg tgg 624
Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp
195 200 205

gga agt ttt tac atg ctg aga aca atg gtt atg gat aaa gtt tgt tgg 672
Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp
210 215 220

ttg aaa cat gta atg tta gat cct gaa aca ggt ctg gac cca ccg aac 720
Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn
225 230 235 240

ttt acg cta cgt gca gca cag ggc ttc gaa tca act gat tat ttc atc 768
Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile
245 250 255

gca ggg tat tgg att tgg aac aaa gtt ttc caa aat ctg gga gta att 816
Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile
260 265 270

ggc tat gaa ccc aat aaa atg acg agt gct gcg tat gat tgg agg ctt 864
Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu
275 280 285

gca tat tta gat cta gaa aga cgc gat agg tac ttt acg aag cta aag 912
Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys
290 295 300

gaa caa atc gaa ctg ttt cat caa tgg agt ggt gaa aaa gtt tgt tta 960
Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu
305 310 315 320

att gga cat tct atg ggt tct cag att atc ttt tac ttt atg aaa tgg 1008
Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp
325 330 335

gtc gag gct gaa ggc cct ctt tac ggt aat ggt ggt cgt ggc tgg gtt 1056
Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val
340 345 350

aac gaa cac ata gat tca ttc att aat gca gca ggg acg ctt ctg ggc 1104
Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly
355 360 365

gct cca aag gca gtt cca gct cta att agt ggt gaa atg aaa gat acc 1152
Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr
370 375 380

att caa tta aat acg tta gcc atg tat ggt ttg gaa aag ttc ttc tca 1200
Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser
385 390 395 400

aga att gag aga gta aaa atg tta caa acg tgg ggt ggt ata cca tca 1248
Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser
405 410 415

atg cta cca aag gga gaa gag gtc att tgg ggg gat atg aag tca tct 1296
Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser
420 425 430

```

```

tca gag gat gca ttg aat aac aac act gac aca tac ggc aat ttc att 1344
Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile
      435                      440                      445

cga ttt gaa agg aat acg agc gat gct ttc aac aaa aat ttg aca atg 1392
Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met
      450                      455                      460

aaa gac gcc att aac atg aca tta tcg ata tca cct gaa tgg ctc caa 1440
Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln
      465                      470                      475                      480

aga aga gta cat gag cag tac tcg ttc ggc tat tcc aag aat gaa gaa 1488
Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu
      485                      490                      495

gag tta aga aaa aat gag cta cac cac aag cac tgg tcg aat cca atg 1536
Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met
      500                      505                      510

gaa gta cca ctt cca gaa gct ccc cac atg aaa atc tat tgt ata tac 1584
Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr
      515                      520                      525

ggg gtg aac aac cca act gaa agg gca tat gta tat aag gaa gag gat 1632
Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp
      530                      535                      540

gac tcc tct gct ctg aat ttg acc atc gac tac gaa agc aag caa cct 1680
Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro
      545                      550                      555                      560

gta ttc ctc acc gag ggg gac gga acc gtt ccg ctc gtg gcg cat tca 1728
Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser
      565                      570                      575

atg tgt cac aaa tgg gcc cag ggt gct tca ccg tac aac cct gcc gga 1776
Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly
      580                      585                      590

att aac gtt act att gtg gaa atg aaa cac cag cca gat cga ttc gat 1824
Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp
      595                      600                      605

ata cgt ggt gga gca aaa agc gcc gaa cac gta gac atc ctc ggc agc 1872
Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser
      610                      615                      620

gcg gag ttg aac gat tac atc ttg aaa att gca agc ggt aat gcc gat 1920
Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp
      625                      630                      635                      640

ctc gtc gag cca cgc caa ttg tct aat ttg agc cag tgg gtc tcc cag 1968
Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln
      645                      650                      655

atg ccc ttc cca arg taa 1986
Met Pro Phe Pro Met
      660

```

<210> 20  
<211> 661  
<212> PRT  
<213> *Saccharomyces cerevisiae*

<400> 20  
Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser  
1 5 10 15  
Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg  
20 25 30  
Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly  
35 40 45  
Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg  
50 55 60  
Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu  
65 70 75 80  
Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe  
85 90 95  
Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe  
100 105 110  
Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val  
115 120 125  
Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn  
130 135 140  
Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly  
145 150 155 160  
Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val  
165 170 175  
Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile  
180 185 190  
Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp  
195 200 205  
Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp  
210 215 220  
Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn  
225 230 235 240  
Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile  
245 250 255  
Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile  
260 265 270  
Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu  
275 280 285

Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys  
290 295 300

Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu  
305 310 315 320

Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp  
325 330 335

Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val  
340 345 350

Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly  
355 360 365

Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr  
370 375 380

Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser  
385 390 395 400

Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser  
405 410 415

Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser  
420 425 430

Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile  
435 440 445

Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met  
450 455 460

Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln  
465 470 475 480

Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu  
485 490 495

Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met  
500 505 510

Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr  
515 520 525

Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp  
530 535 540

Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro  
545 550 555 560

Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser  
565 570 575

Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly  
580 585 590

Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp  
595 600 605

Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser  
610 615 620

Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp  
625 630 635 640

Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln  
645 650 655

Met Pro Phe Pro Met  
660

<210> 21  
<211> 1986  
<212> DNA  
<213> *Saccharomyces cerevisiae*

<220>  
<221> CDS  
<222> (1)..(1983)

<400> 21  
atg ggc aca ctg ttt cga aga aat gtc cag aac caa aag agt gat tct 48  
Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser  
1 5 10 15  
gat gaa aac aat aaa ggg ggt tct gtt cat aac aag cga gag agc aga 96  
Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg  
20 25 30  
aac cac att cat cat caa cag gga tta ggc cat aag aga aga agg ggt 144  
Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly  
35 40 45  
att agt ggc agt gca aaa aga aat gag cgt ggc aaa gat ttc gac agg 192  
Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg  
50 55 60  
aaa aga gac ggg aac ggt aga aaa cgt tgg aga gat tcc aga aga ctg 240  
Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu  
65 70 75 80  
att ttc att ctt ggt gca ttc tta ggt gta ctt ttg ccg ttt agc ttt 288  
Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe  
85 90 95  
ggc gct tat cat gtt cat aat agc gat agc gac ttg ttt gac aac ttt 336  
Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe  
100 105 110  
gta aat ttt gat tca ctt aaa gtg tat ttg gat gat tgg aaa gat gtt 384  
Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val  
115 120 125  
ctc cca caa ggt ata agt tcg ttt att gat gat att cag gct ggt aac 432  
Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn  
130 135 140  
tac tcc aca tct tct tta gat gat ctc agt gaa aat ttt gcc gtt ggt 480  
Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly  
145 150 155 160  
aaa caa ctc tta cgt gat tat aat atc gag gcc aaa cat cct gtt gta 528  
Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val  
165 170 175  
atg gtc cct ggt gtc att tct acg gga att gaa agc tgg gga gtt att 576  
Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile  
180 185 190

gga Gly	gac Asp	gat Asp 195	gag Glu	tgc Cys	gat Asp	agt Ser	tct Ser 200	gcg Ala	cat His	ttt Phe	cgt Arg	aaa Lys 205	cgg Arg	ctg Leu	tgg Trp	624
gga Gly	agt Ser 210	ttt Phe	tac Tyr	atg Met	ctg Leu	aga Arg 215	aca Thr	atg Met	gtt Val	atg Met 220	gat Asp	aaa Lys	gtt Val	tgt Cys	tgg Trp	672
ttg Leu 225	aaa Lys	cat His	gta Val	atg Met 230	tta Leu	gat Asp	cct Pro	gaa Glu	aca Thr	ggg Gly 235	ctg Leu	gac Asp	cca Pro	ccg Pro	aac Asn 240	720
ttt Phe	acg Thr	cta Leu	cgt Arg	gca Ala 245	gca Ala	cag Gln	ggc Gly	ttc Phe	gaa Glu 250	tca Ser	act Thr	gat Asp	tat Tyr	ttc Phe 255	atc Ile	768
gca Ala	ggg Gly	tat Tyr 260	tgg Trp	att Ile	tgg Trp	aac Asn	aaa Lys	gtt Val 265	ttc Phe	caa Gln	aat Asn	ctg Leu	gga Gly 270	gta Val	att Ile	816
ggc Gly	tat Tyr 275	gaa Glu	ccc Pro	aat Asn	aaa Lys	atg Met	acg Thr 280	agt Ser	gct Ala	gcg Ala	tat Tyr	gat Asp 285	tgg Trp	agg Arg	ctt Leu	864
gca Ala	tat Tyr 290	tta Leu	gat Asp	cta Leu	gaa Glu	aga Arg 295	cgc Arg	gat Asp	agg Arg	tac Tyr	ttt Phe 300	acg Thr	aag Lys	cta Leu	aag Lys	912
gaa Glu 305	caa Gln	atc Ile	gaa Glu	ctg Leu	ttt Phe 310	cat His	caa Gln	ttg Leu	agt Ser	ggg Gly 315	gaa Glu	aaa Lys	gtt Val	tgt Cys	tta Leu 320	960
att Ile	gga Gly	cat His	tct Ser	atg Met 325	ggg Gly	tct Ser	cag Gln	att Ile	atc Ile 330	ttt Phe	tac Tyr	ttt Phe	atg Met	aaa Lys 335	tgg Trp	1008
gtc Val	gag Glu	gct Ala	gaa Glu	ggc Gly	cct Pro	ctt Leu	tac Tyr	ggg Gly 345	aat Asn	ggg Gly	ggg Gly	cgt Arg	ggc Gly 350	tgg Trp	gtt Val	1056
aac Asn	gaa Glu	cac His	ata Ile	gat Asp	tca Ser	ttc Phe	att Ile 360	aat Asn	gca Ala	gca Ala	ggg Gly	acg Thr 365	ctt Leu	ctg Leu	ggc Gly	1104
gct Ala	cca Pro	aag Lys	gca Ala	gtt Val	cca Pro	gct Ala	cta Leu	att Ile	agt Ser	ggg Gly	gaa Glu	atg Met 380	aaa Lys	gat Asp	acc Thr	1152
att Ile 385	caa Gln	tta Leu	aat Asn	acg Thr	tta Leu	gcc Ala	atg Met	tat Tyr	ggg Gly	ttg Leu 395	gaa Glu	aag Lys	ttc Phe	ttc Phe	tca Ser 400	1200
aga Arg	att Ile	gag Glu	aga Arg	gta Val 405	aaa Lys	atg Met	tta Leu	caa Gln	acg Thr 410	tgg Trp	ggg Gly	ggg Gly	ata Ile	cca Pro	tca Ser 415	1248
atg Met	cta Leu	cca Pro	aag Lys 420	gga Gly	gaa Glu	gag Glu	gtc Val	att Ile 425	tgg Trp	ggg Gly	gat Asp	atg Met	aag Lys	tca Ser	tct Ser	1296

tca gag gat gca ttg aat aac aac act gac aca tac ggc aat ttc att 1344  
 Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile  
 435 440 445

cga ttt gaa agg aat acg agc gat gct ttc aac aaa aat ttg aca atg 1392  
 Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met  
 450 455 460

aaa gac gcc att aac atg aca tta tcg ata tca cct gaa tgg ctc caa 1440  
 Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln  
 465 470 475 480

aga aga gta cat gag cag tac tcg ttc ggc tat tcc aag aat gaa gaa 1488  
 Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu  
 485 490 495

gag tta aga aaa aat gag cta cac cac aag cac tgg tcg aat cca atg 1536  
 Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met  
 500 505 510

gaa gta cca ctt cca gaa gct ccc cac atg aaa atc tat tgt ata tac 1584  
 Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr  
 515 520 525

ggg gtg aac aac cca act gaa agg gca tat gta tat aag gaa gag gat 1632  
 Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp  
 530 535 540

gac tcc tct gct ctc aat ttg acc atc gac tac gaa agc aag caa cct 1680  
 Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro  
 545 550 555 560

gta ttc ctc acc gag ggg gac gga acc gtt ccg ctc gtg gcg cat tca 1728  
 Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser  
 565 570 575

atg tgt cac aaa tgg gcc cag ggt gct tca ccg tac aac cct gcc gga 1776  
 Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly  
 580 585 590

act aac gtt act att gtg gaa atg aaa cac cag cca gat cga ttt gat 1824  
 Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp  
 595 600 605

ata cgt ggt gga gca aaa agc gcc gaa cac gta gac atc ctc ggc agc 1872  
 Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser  
 610 615 620

gcg gag ttg aac gat tac atc ttg aaa att gca agc ggt aat ggc gat 1920  
 Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp  
 625 630 635 640

ctc gtc gag cca cgc caa ttg tct aat ttg agc cag tgg gtt tcc cag 1968  
 Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln  
 645 650 655

atg ccc ttc cca atg taa 1986  
 Met Pro Phe Pro Met  
 660



&lt;210&gt; 22

&lt;211&gt; 661

&lt;212&gt; PRT

<213> *Saccharomyces cerevisiae*

&lt;400&gt; 22

Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser  
 1 5 10 15

Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg  
 20 25 30

Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly  
 35 40 45

Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg  
 50 55 60

Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu  
 65 70 75 80

Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe  
 85 90 95

Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe  
 100 105 110

Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val  
 115 120 125

Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn  
 130 135 140

Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly  
 145 150 155 160

Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val  
 165 170 175

Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile  
 180 185 190

Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp  
 195 200 205

Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp  
 210 215 220

Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn  
 225 230 235 240

Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile  
 245 250 255

Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile  
 260 265 270

Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu  
 275 280 285

Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys  
290 295 300

Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu  
305 310 315 320

Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp  
325 330 335

Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val  
340 345 350

Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly  
355 360 365

Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr  
370 375 380

Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser  
385 390 395 400

Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser  
405 410 415

Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser  
420 425 430

Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile  
435 440 445

Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met  
450 455 460

Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln  
465 470 475 480

Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu  
485 490 495

Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met  
500 505 510

Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr  
515 520 525

Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp  
530 535 540

Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro  
545 550 555 560

Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser  
565 570 575

Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly  
580 585 590

Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp  
595 600 605

Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser  
610 615 620

Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp  
625 630 635 640

Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln  
645 650 655

Met Pro Phe Pro Met  
660

<210> 23  
<211> 2312  
<212> genomic DNA  
<213> Schizosaccharomyces pombe

<400> 23  
atggcgctctt ccaagaagag caaaactcat aagaaaaaga aagaagtcaa atctcctatc 60  
gacttaccaa attcaaagaa accaactcgc gcttttgagt agcaaccttc agcgctccgaa 120  
acacaatcrg tttcaaataa atcaagaaaa tctaaatttg gaaaaagatt gaattttata 180  
ttgggcgcta ttttgggaat atgcggtgct ttttttttcg ctggttgaga cgacaatgct 240  
gttttcgacc ctgctacggt agataaattt gggaatatgc taggctcttc agacttggtt 300  
gatgacatta aaggatattc atcttataat gtgtttaagg atgcaccttt tactacggac 360  
aagccttcgc agtctcctag cggaaatgaa gtccaagtgt gtcttgatat gtacaatgag 420  
ggatatcgaa gtgaccatcc tgttattatg gtctctggtg ttatcagctc aggattagaa 480  
agttggtcgt ttaataattg ctcgattcct tacttttaga aacgtctttg gggtagctgg 540  
tctatgctga aggcaatggt ccttgacaag caatgctggc ttgaacattt aatgcttgat 600  
aaaaaaacgg gcttggatcc gaagggaatt aagctgcgag cagctcaggg gtttgaagca 660  
gctgattttt ttatcacggg ctattggatt tggagtaaa taattgaaaa ccttgctgca 720  
attggttatg agcctaataa catgttaagt gcttcttacg atggcggtt atcatatgca 780  
aacttagagg aacgtgataa atatttttca aagttaaaaa tgttcattga gtacagcaac 840  
attgtacata agaaaaagg agtggtgatt tctcactcca tgggttcaca ggtracgtac 900  
tattttttta agtgggttga agctgagggc tacggaaaat gtggaccgac ttgggttaat 960  
gatcatattg aagcatttat aaatgtgagt ctcgatggtt gtttgactac gtttctaact 1020  
tttgaataga tatcggggtc tttgatttga gcacccaaaa cagtggcagc gcttttatcg 1080  
ggtagaatga aagctacagg tatgttaatt acattaaaaa tgttaataat taatttttgc 1140  
taaccgtttt aagctcaatt gaatcagttt tgggtctatg ggtaagcaat aaattgttga 1200  
gatttgttac taatttactg tttagtttgg aaaaaatttt tccccgttct gaggtatatt 1260  
caaaaataca aatgtgctct actttttcta acttttaata gagagccatg atggttcgca 1320  
ctatgggagg agttagttct atgcttccca aaggaggcga tgttgatagg ggaatgcca 1380  
gttgggttaag aaatatgtgc tgttaatttt ttattaatat ttaggctcca gatgatctta 1440  
atcaaacaaa tttttccaat ggtgcaatta ttogatatag agaagacatt gataaggacc 1500  
acgatgaatt tgacatagat gatgcattac aattttttaa aaatgttaca gatgacgatt 1560  
ttaaagtcac gctagcgaaa aatttatccc acgggtcttg ttggactgaa aaagaagtgt 1620  
taaaaaataa cgaaatgcg tctaaatgga taaatccgct agaagtaaga acattaaagt 1680  
tactaaatta tactaaccca aatagactag tcttccctat gctcctgata tgaaaattta 1740  
ttgcgttcac ggggtcggaa aaccaactga gagagggtat tattatacta ataactctga 1800  
ggggcaacct gtcaatgatt cctcggttaa tgatggaaca aaagttagaa atgtgagaga 1860  
atttargttt caaacattct attaaactgt ttattagggt attgttatgg atgatggtga 1920  
tggaaattta ccaatattag ccttgggttt ggtggtcaat aaagttaggc aaacaaaaag 1980  
gtttaatcct gctaatacaa gtatcacaaa ttatgaaatc aagcatgaac ctgctgcgtt 2040  
tgatctgaga ggaggacctc gctcggcaga acacgtcgat atacttgga actcagagct 2100  
aaatgtatgt tcattttacc ttacaaattt ctattactaa ctcttgaaat aaggaaatta 2160  
ttttaaaagt ttcatcaggc catggtgact cggtaacaaa ccgttatata tcagatatcc 2220  
agtacggaca taagttttgt agattgcaat taactaacta accgaacagg gaaataataa 2280  
atgagataaa tctcgataaa cctagaaact aa 2312

<210> 24  
<211> 3685  
<212> genomicDNA  
<213> Arabidopsis thaliana

<400> 24  
atgcccttta ttcatecgga aaagccgacg gagaaacat cgacgcgcgc atctgaagag 60  
gtgggtgcacg atgaggattc gcaaaagaaa ccacacgaat ctcccaaatc ccaccataag 120  
aaatcgaaacg gaggagggaat gtgggtcggtc atcgattctt gtgtgtgggt cattgggtgt 180  
gtgtgtgttaa cctgggtggtt tcttctcttc ctttacaacg caatgcctgc gagcttccct 240  
cagtatgttaa cggagcgaat cacgggtcct ttgctgacc cgcgcgggtc taagctcaa 300  
aaagaagggtc ttaaggcgaa acatcctgtt gtcttcattc ctgggactgt caccgggtggg 360  
ctcgagctttt gggaaaggcaa acaatgcgct gatggtttat ttagaaaaacg ttgtgtgggt 420  
ggaacttttg gtgaagtcta caaaagggtg gctcaacaat tctcactctt cctttatatt 480  
gggattttgga ttggatctga tgagatcacg cacttgttgc tcttcaaca tcaactcaac 540  
tttaattcca tgtttgtctg tcttactctt tactttttt ttttttgat gtgaaacgct 600  
attttcttaa gagactattt ctgtatgtgt aaggtaagcg ttccaaggac gtaattggct 660  
tggactattt ctggttgatt gtttaactta ggatataaaa tacttgctt ggaatttcaa 720  
gtcactttat tgcctaatct gttgctagac atgcctaga gtccgttcat aacaagttac 780  
ttcctttact gtctgtgct gtatatttag ctttgtgtag cgtataatga agtagtgtt 840  
tatgttttgt tgggaataga gaagttctaa ctacatctgt ggaaagtgtg ttcaggctgt 900  
gatagaggac tgttgcctta ttatccaact atgtatatgt gtaattaaag ctagtctctt 960  
tttgatcttt cagctcaatg tgcctttctc aactttttt tcaatttcaa agtttccat 1020  
cgagttttat cacatgtctt gaattctgtc cactctctgt ctgttatcca gctttgaact 1080  
ctcccgacc ctgctatgga tatattaaaa aaaaagtgtt ttgtgggttg catctttgtt 1140  
acgacttgca tcttctctt tgggtcagc gtccatgttt ttgctatggt agagatgggc 1200  
aatgttattg ttgatgttaa cagtgttata gttgatagta tcttaactaa tcaatttat 1260  
cttcgattca ggcctctatg ttgggtggaa cacatgtcac ttgacaatga aactgggttg 1320  
gatccagctg gtattagagt cagagctgta tcaggactcg ttgctgctga ctactttgtc 1380  
ctcgctact ttgtctggc agtgcctgatt gctaaccttg cacatactgg atatgaagag 1440  
aaaaatatgt acatggctgc atatgactgg cggctttcgt ttccagaacac agaggttctt 1500  
ttctcatcgt tcttctatt attctgttc atgttacgtt tcttcttca ttacttaagg 1560  
cttaaatatg ttcctgtgt aattaatagc tacgtgacca gactcttagc cgtatgaaaa 1620  
gtaatataga gttgatggtt tctaccaacg gtggaaaaaa agcagttata gttccgcat 1680  
ccatgggggt cttgtatttt ctacatttta tgaagtgggt tgaggcacca gctcctctgg 1740  
gtggcggggg tggccagat tgggtgtgcaa agtatattaa ggcgggtgat aacattggtg 1800  
gaccatttct tgggttcca aaagctgttg cagggtttt ctctgctgaa gcaaggatg 1860  
ttgcagttgc caggtattga acatcctgtt atacttttga tgatcagaac cttggctctg 1920  
gaactcaaaag ttattctact aaatatcaat tctaaacaac ttgctatatt atcgtgcaa 1980  
ctgacattgg ttgattatt ttgctgctta tgttaactgaa actctcttga gattagacaa 2040  
atgatgaatt gataattctt acgcattgct ctgtgatgac cagtttctta gcttcgacga 2100  
taacatttgt catactgct tttgagggc attgaatttt gctatggaaa gcgctggagc 2160  
ttccatgctt gcattcttta ccaattagcg ttattctgct tcttcaatt ttcttgata 2220  
tgcatctatg gtcttttatt tctcttaac taaagactcg ttggattagt tgcctatta 2280  
gtcacttggc tcttaatat agaactttac tttcttcgaa aattgcagag cgattgccc 2340  
aggattctta gacaccgata tatctagact tcagacctg cagcatgtaa tgagaatgac 2400  
acgcacatgg gactcaacaa tctctatgtt accgaaggga ggtgacacga tatggggcgg 2460  
gcttgattgg tcaccggaga aaggccacac ctgttgtggg aaaaagcaaa agaacaacga 2520  
aacttgtgtg gaagcaggtg aaaacggagt ttccaagaaa agtctgtta actatggaag 2580  
gatgatattt tttggaaag aagtagcaga ggctgcgcca tctgagatta ataatactga 2640  
cttctgagta aggacatata aatcataata aacctgtac atttgtgat tgtatgatga 2700  
atatctgcac attttatctg gtgaagggtg ctgtcaagg tcagagtatc ccaaatcaca 2760  
cctgtcgta cgtgtggaca cagtaccatg acatgggaat tgcgtggatc aaagctatcg 2820  
ctgagtataa ggtctacact gctggtgaag ctatagatct actacattat gttgctcta 2880  
agatgatggc gcgtgggccc gctcattctt ctatggaa ttgctgatgat ttggatgaca 2940  
ccaagtatca agatcccaaa tactgggtcaa atccgttaga gacaaagtta gtcattctt 3000  
gattccaact gtatccttgc tcttgatgca ttatcagct ttttgtctc ggtctgtctg 3060  
gatattgttt tcagctcaaa gcttacaagg ctgtttctga gcttttctca aaaagcctcg 3120  
ctcagtaata ttgaggtgct aaagttgata catgtgactc ttgcttataa atctccgct 3180

tggtttgttc tgctttttca gattaccgaa tgctcctgag atggaaatct actcattata 3240  
cggagtgggg ataccaacgg aacgagcata cgtatacaag cttaaccagt ctcccgacag 3300  
ttgcatcccc ttccagatat tcacttctgc tcacgaggag gacgaagata gctgtctgaa 3360  
agcaggagtt tacaatgtgg atggggatga aacagtaccc gtcctaagtg ccgggtacat 3420  
gtgtgcaaaa gcgtggcgtg gcaagacaag attcaaccct tccggaatca agacttatat 3480  
aagagaatac aatcaccttc cgcgggctaa cctgttggaa gggcgcggga cgcagagtgg 3540  
tgcccatggt gatatcatgg gaaactttgc tctgatcgaa gatatcatga gggttgccgc 3600  
cggaggtaac cggctctgata taggacatga ccaggccac tctggcatat ttgaatggtc 3660  
ggagcgtatt gacctgaagc tgtga 3685

<400>	25																	
agaaacagct	ctttgtctct	ctcgactgat	ctaacaatcc	ctaattctgtg	ttctaaattc												60	
ctggacgaga	tttgacaaag	tcggtatagc	ttaacctggt	ttaatttcaa	gtgacagat												119	
atg	ccc	ctt	att	cat	cgg	aaa	aag	ccg	acg	gag	aaa	cca	tcg	acg	ccg			167
Met	Pro	Leu	Ile	His	Arg	Lys	Lys	Pro	Thr	Glu	Lys	Pro	Ser	Thr	Pro			
1				5				10				15						
cca	tct	gaa	gag	gtg	gtg	cac	gat	gag	gat	tcg	caa	aag	aaa	cca	cac			215
Pro	Ser	Glu	Glu	Val	Val	His	Asp	Glu	Asp	Ser	Gln	Lys	Lys	Pro	His			
			20			25						30						
gaa	tct	tcc	aaa	tcc	cac	cat	aag	naa	tcg	aac	gga	gga	ggg	aag	tgg			263
Glu	Ser	Ser	Lys	Ser	His	His	Lys	Xaa	Ser	Asn	Gly	Gly	Gly	Lys	Trp			
			35			40						45						
tcg	tgc	atc	gat	tct	tgt	tgt	tgg	ttc	att	ggg	tgt	gtg	tgt	gta	acc			311
Ser	Cys	Ile	Asp	Ser	Cys	Cys	Trp	Phe	Ile	Gly	Cys	Val	Cys	Val	Thr			
			50			55						60						
tgg	tgg	ttt	ctt	ctc	ttc	ctt	tac	aac	gca	atg	cct	gcg	agc	ttc	cct			359
Trp	Trp	Phe	Leu	Leu	Phe	Leu	Tyr	Asn	Ala	Met	Pro	Ala	Ser	Phe	Pro			
			65			70						75			80			
cag	tat	gta	acg	gag	ccg	aat	cac	gng	tcc	ttt	gcc	tta	ccc	g				402
Gln	Tyr	Val	Thr	Glu	Pro	Asn	His	Xaa	Ser	Phe	Ala	Leu	Pro					
			85			90												

<210> 26  
<211> 643  
<212> cDNA  
<213> Zea mays

<220>  
<221> CDS  
<222> (1)..(402)

<400> 26  
cgg gag aaa ata gct gct ttg aag ggg ggt gtt tac tta gcc gat ggt 48  
Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly  
1 5 10 15  
gat gaa act gtt cca gtt ctt agt gcg ggc tac atg tgt gcg aaa gga 96  
Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly  
20 25 30  
tgg cgt ggc aaa act cgt ttc agc cct gcc ggc agc aag act tac gtc 144  
Trp Arg Gly Lys Thr Arg Phe Ser Pro Ala Gly Ser Lys Thr Tyr Val  
35 40 45  
aga gaa tac agc cat tcg cca ccc tct act ctc ctg gaa ggc agg ggc 192  
Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly  
50 55 60  
acc cag agc ggt gca cat gtt gat ata atg ggg aac ttt gct cta att 240  
Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile  
65 70 75 80  
gag gac gtc atc aga ata gct gct ggg gca acc ggt gag gaa att ggt 288  
Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly  
85 90 95  
ggc gat cag gtt tat tca gat ata ttc aag tgg tca gag aaa atc aaa 336  
Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Lys  
100 105 110  
ttg aaa ttg taa cct atg gga agt taa aga agt gcc gac ccg ttt att 384  
Leu Lys Leu  
115  
gcg ttc caa agt gtc ctg cctgagtgca actctggatt ttgcttaaatt 432  
attgtaattt ttcaogcttc attcgtccct ttgtcaaatt tacatttgac aggacgccaa 492  
tgcgatacga tgttgtaccg ctattttcag cattgtatat taaactgtac aggtgtaagt 552  
tgcatttgcc agctgaaatt gtgtagtcgt tttctttacg atttaatanc aagtggcgga 612  
gcagtgcccc aagcnaaaaa aaaaaaaaaa a 643



&lt;210&gt; 27

&lt;211&gt; 115

&lt;212&gt; PRT

&lt;213&gt; Zea mays

&lt;400&gt; 30

Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly  
1 5 10 15  
Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly  
20 25 30  
Trp Arg Gly Lys Thr Arg Phe Ser Pro Ala Gly Ser Lys Thr Tyr Val  
35 40 45  
Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly  
50 55 60  
Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile  
65 70 75 80  
Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly  
85 90 95  
Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Lys  
100 105 110  
Leu Lys Leu  
115

<210> 28  
<211> 516  
<212> cDNA  
<213> *Neurospora crassa*

<400> 28  
ggtggcgaag acgangggcg aagttggagg ctaacgagaa tgacnccgg agatggatct 60  
accctctaga gacacgacta centtgcacc cagcctcaag gntacngtt tntatgggta 120  
ggaagccgac ggagcgagcc tacatctatc tggcgccga tcccgggacg acaacgcac 180  
tttagatgac gatcgatacg actttgactn aggggcacat tgaccacggg gtgattttgg 240  
gcgaaggcga tggcacagtg aaccttatga gtttggggta cctgtgcaat aaggggtgga 300  
aaatgaagag atacaatcct gcgggctcaa aaataaccgt ggtcgagatg ccgcatgaac 360  
cagaacgggt caatccgaga ggaggccga atacggcgga cttaaâtatg tagaaaagg 420  
tgaaatttat gaagagtaat taaatacggc acataggtta ctcaatagta tgactaatta 480  
aaaaaaaaatt ttttttctaa aaaaaaaaaa aaaaaa 516

<210> 29  
<211> 1562  
<212> genomic DNA  
<213> Arabidopsis thaliana

<400> 29  
atgaaaaaaaa tatcttcaca ttatttcggta gtcatagcga tactcgttgt ggtgacgatg 60  
acctcgatgt gtcaagctgt gggtagcaac gtgtaccctt tgattctggt tccaggaaac 120  
ggaggttaacc agctagaggt acggtcggac agagaatata agccaagtag tgtctgggtg 180  
agcagctggt tatatccgat tcataagaag agtggtagat ggtttaggct atggttcgat 240  
gcagcagtggt tattgtctcc ctccaccagg tgcttcagcg atcgaatgat gttgtactat 300  
gacctgatt tggatgatta ccaaaatgct cctgggtgtc aaaccgggt tccatatttc 360  
ggttcgacca aatcacttct atacctcgac cctcgtctcc ggttagtact tccaagata 420  
tatcattttg ggacatttgc ataatagaac aaatagacat aaatttgggg gattattggt 480  
atatcaatat ccatttatat gctagtcggt aatgtgagt ttatgttagt atagttaatg 540  
tgagtgttat gtgatttctc attttaaatg aagctagaaa gttgtcgttt aataatgttg 600  
ctatgtcatg agaattataa ggacactatg taaatgtagc ttaataataa ggcttgattt 660  
gcagagatgc cacatcttac atggacatt tggtagaagc tctagagaaa aaatgcgggt 720  
atgttaacga ccaaaccatc ctaggagctc catatgattt caggtagggc ctggctgctt 780  
cggggccacc gtcccgtgta gccctacagt tccataaga cctcaaacaa ttggtagaaa 840  
aaactagcag cgagaacgaa ggaagccag tgatactct ccccatagc ctaggaggac 900  
ttttcgtcct ccatttctc aaccgtacca ccccttcctg gcgcgcgaag tacatcaaac 960  
actttgttgc actcgtctgc ccattgggtg ggacgatctc tcagatgaag acatttgctt 1020  
ctggcaacac actcgtgtgc cctttagtta accctttgct ggtagacgg catcagagga 1080  
cctccgagag taaccaatgg ctacttccat ctaccaaagt gtttcacgac agaactaac 1140  
cgcttgcctt aactccccag gtttaactaca cagcttacga gatggatcgg tttttgcag 1200  
acattggatt ctcaacagga gttgtgcctt acaagacaag agtgttgctt ttaacagagg 1260  
agctgatgac tccgggagtg ccagtcactt gcataatag gagaggagt gatcacccgg 1320  
aggttttgat gtatggaaaa ggaggattcg ataagcaacc agagattaa gattggagatg 1380  
gagatgggac ggttaatttg gcgagcttag cagctttgaa agtcgatagc ttgaacaccg 1440  
tagagattga tggagtctcg catacatcta tacttaaaga cgagatcgca cttaaagaga 1500  
ttatgaagca gatttcaatt attaatatg aattagccaa tgtaaatgcc gtcaatgaat 1560  
ga 1562

```
<210> 30
<211> 3896
<212> genomic DNA
<213> Arabidopsis thaliana
```

<400> 30						
atggggagcgga	atttcgaaatc	agtaacggct	tccctcaccg	tcacgcgcgt	ttttttcttg	60
atcttggcgtg	gccgaacttc	ggcggaggat	gagacgcgat	ttcagcgcgga	ctactcgaag	120
ctatcgggga	taactcattcc	gggattttgcg	tcagcgcagc	tacgagcgtg	gtcgatcctt	180
gactgtccat	acactccgtt	ggacttcaat	ccgcctcgacc	tcgtatggct	agacaccact	240
aaggctccgtg	atcttcattt	ccttcgcctcc	tatatctgtc	ggctcgagtca	cttggtgatg	300
aattccaagc	gaatatagac	aatgaagcat	gtcctcgctc	tcttattgtg	tcgtttcata	360
gtcaaacatg	acgctttctga	atcttgagtt	agagctatat	aaaacagctg	actcggcgag	420
tgtttcccat	cgcttttggg	tcgtcaaatg	tagcgcaatg	aattgttaat	tagtctgcgc	480
tttttattca	actagatctg	caagtttttc	agagtgctca	atagtagtca	gaaaatgtta	540
ggtcattttta	cttgtgcatt	gtgattcttt	tggttgtttg	ttactgatcg	acgtgatgga	600
tgtttttacag	cttctttctg	ctgtcaactg	ctggtttaag	tgtatggctg	tagatcctta	660
taatacaaac	gacctaccgc	agtgtaaagc	acggcctgac	agtggtcttt	cagccatcac	720
agaattggat	ccagggttaca	taacaggtag	tttcggattt	tcttttcttt	tgagttctct	780
tcaatttgat	atcatcttgt	tgtgatataa	tatggctaag	ttcattaatt	tgggtcaatt	840
tcaggctcctc	ttctctactgt	ctggaagag	tyggtttaagt	gggtgtgtga	gtttgggtata	900
gaagcgaatg	caatttgtgc	tgttccatag	gattggagat	tgtcaccaac	caaatgtgaa	960
gagcgtgacc	ttatctttca	caagctcaag	ttagtctcta	tcaggctaat	gtctttttatc	1020
ttctcttttt	atgtaagata	agctaagagc	tctggctctc	ttcttttttg	cagggttgacc	1080
tttgaaaactg	cttcaaaaact	cogtggcggc	ccttctatag	tatttgccca	ttcaatgggt	1140
aataatgtct	tcagataact	tctggaatgg	ctgaggcctag	aaattgcacc	aaaacattat	1200
ttgaagtgct	ttgatcagca	tatccatgct	tatttcgctg	ttggtagccg	ctcactatcc	1260
ttgaagttacc	attttatttt	ttctctaat	ggggggagta	tgttgtgact	tactggattg	1320
agctcgatac	ctgattttgtc	gttgattttag	gagctcctct	tcttgggtct	gttgaggcaa	1380
tcaaatctac	tctctctggg	gtaacgtttg	gccttctgt	ttctgagggt	acctctgact	1440
tctcttttagt	tttaagttag	tgatataaac	caggctcttat	aactcactgg	attttctctt	1500
tgaagaatt	acttttgta	attgaactgc	ttagtcgcgt	attggtatct	tagatcttga	1560
agtgcctagtt	atcaagaac	atattgtggg	tagtataact	gtcagcgcc	ttagctaata	1620
caaccaaacc	acatgtacac	tgatttagtt	ttcagattat	tacggtagac	tttaagttga	1680
gaagaaactt	tgaactgaac	ctctttattt	taataggcta	tagtttgtt	atgaaatac	1740
tgtgacatat	tgacatgcgc	ttctcatggt	ttttgtggc	agggcttcag	ggaaactgct	1800
ggttgttgtc	caattctttt	gcgtcgtcat	tgtggcttat	gccatttca	aagaattgca	1860
agggtgataa	cacattctgg	acgcattctt	ctggggctgc	tgcaaagaaa	gataagcgcg	1920
tacaccactg	tgatgaagag	gaatatcaat	caaaaatact	tggtcggcgc	acaaatacta	1980
ttaacataga	aatttcttcc	actagcgggt	agactctgta	tatgcaactg	taacataaac	2040
aaaagtttca	ccaagaatgt	tcactctcat	atttcgttcc	tttgatgtgt	atccatcagt	2100
tacagaaaca	gctctagtc	acatgaccag	catggaatgt	ggccttccca	cccttttgtc	2160
tttcacagcc	cgtgaactag	cagatgggac	tcttttcaaa	gcaatagaag	actatgacc	2220
agatagcaag	aggatgttac	accagttaaa	gaagtacgta	cccttctctg	tgataagaaa	2280
tattgctcat	cgatcatcac	cttgtgcctt	cttgtactgc	aaattgtttt	gttttaactc	2340
ctatacaac	tgtctatag	tttgtgcctt	cttactataa	gaaacaagta	taatacagaa	2400
ccctattatt	gacctcagt	tccctcctta	tattatggaa	tgtctttttc	gtttacagtt	2460
atgaatgcaa	aaggggggat	tttagttgat	tgatttctct	attctctagt	ttgttttgac	2520
taatagcgtc	aatttttgtt	tcctagcaaa	tcttttgtga	ttatatataa	catgctaact	2580
atacctttca	ggttgtatca	tcctgaccct	gtttttcaac	ctctgactcc	ttgggaagaga	2640
ccacctataa	aaaatgtatc	ttgcataat	gggtgctcat	taaagacaga	gggtatgatgc	2700
atctccaata	tcacattatg	cgttgacttc	gtaattatat	tccccatttg	gtttgcataa	2760
tctttctgaa	tcattgatct	tcttctcctt	tgcattctat	gctatcaagc	gttaaaagga	2820
ctaaaagtat	gaagcttgtt	gtcatagggt	ggtttattact	ttgcccacag	tggcaaacct	2880
tactctgata	actggatcat	cacggatatc	atttatgaaa	ctgaagggtc	cctcgtgtca	2940
aggtaatttt	cgcgaatggc	agaagtataa	caggaaggca	aagtcttctg	tatcagttca	3000
gtgcattgtt	acctcagctg	cataagcaaa	ttatatcaac	actaaaattt	aggtactttc	3060
ttatcatctc	ttttgagctt	atgttgatgt	cagtgggctta	aagtggggaag	aggtgtttgca	3120
tgaaacatga	cactctctatc	aaaagataact	agcaaaaacaa	aactaaccca	tttctgaatt	3180

tcataattatt aggagtagtc gtgctttttaa aaaatttggt ttaagaaacc gaaaaactag 3240  
ttcatatctt gattgtgcaa tatctgcagg tctggaactg tgggtgatgg gaacgctgga 3300  
cctataactg gggatgagac ggtaagctca gaagttgggt ttgaaattat cttcttgcaa 3360  
actactgaag actaagataa tacttgcttc tggaacactg cttgctatgt tctctagtac 3420  
actgcaatat tgactctccg ctacttttat tgattatgaa attgatctct tataggtacc 3480  
ctatcattca ctctcttggg gcaagaattg gctcggacct aaagttaaca taacaatggc 3540  
tccccaggta ctcttttcta gtctctcacc ttatatagat caaaccttaa gtgtraccttt 3600  
ctgggttargt gttgatttac ctccaatttg ttctttctaa aaatcatata tctctgtact 3660  
cctcaagaac ttgtattaat ctaaacgaga ttctcattgg gaaaataaaa caacagccag 3720  
aacacgatgg aagcgacgta catgtggaac taaatgttga tcatgagcat gggtcagaca 3780  
tcatagctaa catgacaaaa gcaccaaggg ttaagtacat aaccttttat gaagactctg 3840  
agagcattcc ggggaagaga accgcagtct gggagcttga taaaagtggg tattaa 3896

<210> 31  
<211> 709  
<212> cDNA  
<213> tomato

<400> 31  
ctggggccaa aagtgaacat aacaaggaca ccacagtcag agcatgatgt ccagatgtac 60  
aagtgcacat aaatatagag catcaacatg gtgaagatat cattcccaat atgacaaagt 120  
tacctacaat gaagtacata acctattatg aggattctga aagttttcca gggacaagaa 180  
cagcagtttg ggagcttgat aaagcaaata acaggaacat tgtcagatct ccagctttga 240  
tgcgggagct gtggcttgag atgtggcatg atattcatcc tgataaaaag tccaagtttg 300  
ttacaaaagg tgggtgtctga tctcactat tttcttctat aaatgtttga gtttgtattg 360  
acattgtaag tattgcaaca aaaagcaaag cgtgggcctc tgaggggatga ggactgctat 420  
tggtgattac ggaaagctcg atgtgcatgg gctgaacatt gtgaatacag gttagaatat 480  
tcaaattata ttttgcaaaa tttctctctt ttgtgtattt aggccacctt tccccggtca 540  
caacgatgca gatatgtatt cggggatgtt cacctgggac agagttgcag attgaagagt 600  
tctacatctc acatcctgtc acactatgtg tgatatttaa gaaactttgt ttggcggaac 660  
aaceagtttg cacaaacatt tgaagaagaa agcgaaatga ttcagagag 709